

|CURIEUX|

ACADEMIC JOURNAL

March Issue

Part 1 Issue 48

Editing Staff

Chief-In-Editor

Caroline Xue

George Hasnah

Chief of Operations

Anshal Vyas

Assisting Editors

Olivia Li

Luzia Thomas

Madalyn Ramirez

Shefali Awasthi

Soumya Rai

Sarah Strick

Table Of Contents

Page 6: The Hindrance of Corporate Employee Productivity: Exploring Toxic Workplace Culture By Eric Tam

Page 14: Examining the Reading Brain By Jasmine Ndula

Page 20: Politics and Paternalism: The Conditions Behind Capabilities By Matthew Jeong

Page 32: MBRP as an Effective Method to Treat Substance Use Disorders By Arul Muthuvelu

Page 39: Trends of Down Syndrome Patients Developing Leukemia: A Review of Risk Factors and Early Cancer Detection Using Liquid Biopsy By Tanisha Arora

Page 47: The Utility of Art in Depicting the Contrast Between the Women of Sparta and Athens By Vivienne Tang

Page 68: Is Chemosynthetic Technology Viable and Realistic to Use in the Context of How Much Energy It Can Produce and its Environmental Impact? By Harika Chodavarapu

Page 77: Predictive Modeling of Stock Market Prices Using Machine Learning By Aidan Lightman

Page 84: The Cardiovascular Implications of Mutated Genes in Breast Cancer By Sai Saharsh Gumudavelly

Page 97: Shifts in Migration Arrival Timing of Understudied South American Tern (*Sterna hirundinacea*) in the Falkland (Malvinas) Islands in Response to Climate Change By Daphne Zhu

Page 107: Exploring Exoplanet Habitability: Equilibrium Temperature, Observational Constraints, and Implications for the Habitable Zone By Ngoc Ha Bao Le¹ (Student) and Ana-Maria Piso-Grigore² (Mentor)

Page 117: A Review On the Effectiveness of Albumin Nanocarrier Delivery Vehicles When Combined With Paclitaxel in the Treatment of Metastatic Breast Cancer By Robin Chai

Page 129: From Lev to Euro: Prospective Benefits of Eurozone Integration for Bulgaria By Maria Kamenova

Page 141: Comparing Classical Music Styles Using Markov Chains By Ashita Garg

Page 149: Building the Bubble: China's Vulnerability Due to Rapid Economic Growth Fueled by Speculative Real Estate By Hans Zhu

Page 159: Bridging the Divide: Comparative Perspectives on Women's Empowerment in Developed and Developing Nations By Syed Zahir Hussain

Page 175: Math and Music: Polyrhythms By Moksh Naik

Page 181: BMI's Impact on Ability to Hold Front Lever Holds By London Bruni

Page 188: The Neural Basis of Emotion Regulation in Major Depressive Disorder: Insights from Cognitive Behavioral Therapy and Neuroimaging Studies By Xuehan (Helen) Ye

Page 204: Innovations for *In Vivo* Gene Therapy Delivery and the AI-Driven Future of Genetic Medicine By Norman Lwin

Page 225: Impacts of Global Events on the Consumer Economy of the UAE By Arnav Malhotra

Page: 237: AI in the Creative Music Economy: Legal Frameworks and Small Artists By Carly Ding

Page 246: The Future of Huntington's Disease Treatments Through Genetic Editing By Adithi Chakkerla

Page 252: How can Neurocircuitry Help to Improve the Understanding and Treatment of Depression? By Alexis Lavelly

Page 272: Differences in Autism Diagnosis Rates Due to Various Racial, Geographic, Genetic, and Biological Factors By Divya Agarwal

Page 287: The Ugly Truth Behind the Largest Enterprises: How Are Small Businesses (SMEs) Being Impacted? By Samanvi Hublikar

Page 293: Optimal Tilt Angle For Solar Collectors In Telangana, India By Varun Ponnala

Page 301: Choking Civic Spaces: Digital Censorship and Its Effects on Indonesian Civil Society Organizations By Jayden Christian Budiman

Page 309: From Concealment to Confrontation: France's Shifting Memory of Nazi Occupation and the Algerian War By Jean-Louis Andreani

Page 327: The Effects of Physical Activity and Sleep on Stress and Happiness: A Comparative Study of General and International Alternative School Students By Yoseop Jeong

Page 340: Repairing the Cracked Foundation of Financial Equality: Policy Solutions to Reverse the Legacy of Housing Discrimination By Anthony Borrelli

Page 352: Pros and Cons of Digital Media Use as a Coping Strategy By Arjun Chopra

Page 365: An Analysis of Battery Swap Technology for Electric Vehicles By Xiang Jiang

Page 374: Navigating the Nexus: Tourism, Sustainability, and Homelessness in Hawaii's Changing Landscape By Justin Chun

Page 382: Effects of Monetary Sanctions on the Intergenerational Financial Stability of Families with Incarcerated Members By Hyeju Yoon

Page 395: The Potential Use of Combined High-Frequency Repeated Transcranial Magnetic Stimulation (HF-rTMS) and Virtual Reality (VR) for Post-Stroke Rehabilitation – A Literature Review By Harini Prabhu Shankar

Page 411: Using Algae as A Catalyst for Carbon Neutrality By Amber Dogra

Page 420: The Relationship Between Green Infrastructure and Carbon Emissions: A Literature Review By Srinath R. Dadireddy (student)¹

Page 431: To What Extent Does Public Opinion Impact Decisions Made in the Judicial System of The USA and the UK? By Shoni S.N. Doshi^{1*} Emmah K.S. Wabuke²

Page 446: Navigating the Digital Shift: The Impact of E-Commerce on Brick-and-Mortar Retail Strategies and Consumer Behavior By Rebecca Blumberg

Page 455: Market Analysis of Alibaba Group By Bryan Li

Page 463: Sociocultural Norms Shaping Sports Participation Among South Asian & Arab Female Athletes By Devika Hajarnavis

Page 477: Constraining Observables in Neutron Stars with Varying Equations of State By Sahil Arora

Page 495: Project NeoWave: A Comparative Analysis of Conservatism and Gender Dynamics Impacting Women and Families in the USA and India Since 2010, with Policy Recommendations By Shivansh Gupta

Page 516: The Microchip Revolution: Jack Kilby's Invention and the Past, Present and Future of Technology By Alexander Recce

Page 524: Bridging Scientific Disciplines & Cultures: A Multidisciplinary and Multicultural Approach to Palliative Care By Hannah Yin

Page 548: Tariffs and Their Unexpected Consequences: The Impact of US-China Trade Restrictions on the US and Third-Party Countries By Ian Wong

The Hindrance of Corporate Employee Productivity: Exploring Toxic Workplace Culture

By Eric Tam

Abstract

This study investigates the impact of toxic workplace culture on corporate employee productivity, emphasising the primary factors of toxic leadership and neglect of employee well-being. A review of over 50 pieces of existing literature from scholarly sources was conducted with a focus on key themes such as toxic workplace culture, employee productivity & well-being, interpersonal conflict, and impact on mental health. This review reveals that toxic leadership, characterised by a lack of support, poor communication, and abusive behaviours, undermines employee well-being. Similarly, the lack of consideration of employees' well-being leads to increased dissatisfaction, burnout, and decreased productivity among employees. The findings show that employees are more likely to resign due to a toxic environment than inadequate compensation, highlighting the need for organisations to address these issues. By implementing effective leadership training focused on trust-building and open communication as well as support systems for mental health, organisations can create a healthy workplace culture. Furthermore, encouraging inclusivity and celebrating diversity can mitigate interpersonal conflicts arising from cultural differences, which can enhance employee morale and productivity. This study emphasises the necessity for coordinated efforts from all management levels to cultivate a respectful and supportive workplace that prioritises employee well-being.

Introduction

In the current era of globalisation, the interference on corporate employee productivity due to toxic workplace culture has become a significant concern when considering the effects on individual employees and workplaces as a whole (Rasool et al., 2021). Corporate employee productivity, defined by Coker (2011) as *"the level of employees' performance in relation to attendance, work quality, the capacity of performance, and personal factors,"* encompasses a multitude of factors beyond mere efficiency. This kind of productivity is often hampered by a toxic workplace culture, where Chamberlain and Hodson (2010) characterise these environments as having *"high levels of interpersonal conflict, a lack of worker autonomy, and a high level of disorganization."* The prevalence of toxic workplace culture is reinforced by the *American Psychological Association's* 2024 Work in America workforce survey, where 15% of respondents labelled their workplace as toxic (Sleek, 2024). Linking the impacts of a toxic workplace on employee productivity, a statistic given by *Financial Post* states that companies in Canada lost 40 billion dollars because of decreased employee productivity in 2022 (Wells, 2022). Due to the relationship between toxic workplace culture and decreased employee productivity, companies are beginning to see the rippling effect of lower employee productivity as a whole (Sull & Sull, 2022). The widespread impact and significance of toxic workplace culture affecting employee productivity in the workplace has led to the research question: to what extent is employee productivity affected by a toxic workplace culture? To answer this question, a comprehensive

review and synthesis of recent and relevant sources were conducted, with a focus on identifying key themes of toxic workplace culture, employee productivity and well-being, interpersonal conflict, and impact on mental health to inform the findings in this article. The literature review encompassed a diverse range of scholarly sources, including peer-reviewed articles from reputable academic journals, research reports from professional organisations, government and institutional reports that offer data and statistics, and credible news sources reporting on workplace studies and trends. This approach to selecting sources ensures a well-rounded study on toxic workplace culture and its effects on corporate employee productivity, drawing from various perspectives.

Factors Affecting Corporate Employee Productivity

A review of over 50 literature suggests that corporate employee productivity is hindered by a toxic workplace, which is created through two major factors: toxic leadership and the limited consideration of individual employee well-being. Toxic leadership is characterised by a lack of support and trust, poor communication, and abusive behaviours, which undermine employee morale (Sull & Sull, 2022; Gandolfi & Stone, 2022; Peetz, 2019). On the other hand, the neglect of employee well-being is where supervisors fail to prioritise mental health and work-life balance, leading to dissatisfaction, burnout, and decreased productivity (Anjum et al., 2018; H.H.S., 2024; Sleek, 2024).

Toxic Leadership

Before looking at the effects that toxic workplace culture has on employee productivity, it is important to diagnose the factors that drive these harmful cultures. Sull and Sull (2022) executed a study and identified that the three predominant factors are toxic leadership, toxic social norms, and poor work design. While these are the most common factors that drive a toxic workplace culture, toxic leadership emerged as the best predictor of toxic culture. Toxic leadership is characterised by behaviours such as micromanagement, abusive language, and self-centeredness, which create a hostile and negative work environment (Gandolfi & Stone, 2022). These behaviours are indicative of a leadership style that fundamentally lacks trust in employees. Therefore, toxic leadership can be associated with a management style without trust, where Peetz (2019) states that “*low trust management styles are often characterized by low worker autonomy and close supervision.*” Essentially, it can be said that toxic leadership stems from low levels of trust, which is often marked by micromanagement. Furthermore, Healy and Dowell (2017) explain that supervisor fears can be a core reason for poor management, detailing that said fears can lead to being hypercritical of errors while simultaneously ignoring accomplishments. Healy and Dowell (2017) further suggest that poor management can lead to lower levels of productivity due to an increase in dissatisfaction. This highlights how mismanagement, a characteristic of toxic leadership, can decrease productivity in the workplace (Healy & Dowell, 2017). Connecting these pieces of evidence, it can be summarised that toxic

leadership is the main driver of toxic workplace culture, directly impacting employee productivity.

While these sources discuss how leadership style is a determining factor of toxic culture, Thomas and Chaleff (2017) propose a counter-argument by placing responsibility into the hands of the employees to mitigate toxicity in the workplace. Thomas and Chaleff (2017) state that *“those who are just obeying orders and conforming to the culture are just as culpable as those giving the orders, and that individuals need to stand against immoral, unethical, or illegal behavior.”* However, this perspective overlooks a critical barrier: fear. Shoss, (quoted by Sleek, 2023) emphasises this by stating that *“toxic workplaces drain all the energy and excitement out of employees and replaces it with fear.”* It is unjust to say that employees are “just as culpable as those giving the orders,” as Shoss asserts that toxic workplaces instil fear in employees, making it extremely difficult to speak up. Healy and Dowell reinforce Shoss’ claim by stating that *“employees fear that they will lose their job, lose status within the organization, lose wage increases, or lose worker autonomy”* (Healy & Dowell, 2017, p. 29). In corporate settings with clear hierarchical systems, employees are vulnerable to toxicity in the workplace, due to the power dynamic that can cause feelings of inferiority, making it hard to stand against injustice (Looman et al., 2022). Therefore, due to the power dynamic that exists between employees and their respective supervisors, employees cannot be held equally accountable for feeding into toxic workplace culture, which is why the argument by Thomas and Chaleff falls.

Effects on Well-being

Proceeding on, employee productivity is lowered due to the effects of toxic workplace culture, where there is limited consideration of individual employee well-being. Specifically, there are significant impacts on mental health, job productivity, and interpersonal conflict. A study conducted by MIT Sloan analysed over 34 million online employee profiles, producing a statistic that employees are 10.4 times more likely to resign due to toxic workplace culture than compensation (Sull et al., 2022). This statistic goes to show that toxicity in the workplace is the prevailing reason as to why employees resign, as it has the most impact. Additionally, Sull and Sull (2022) state that workers who experience the elements of a toxic culture are more likely to suffer from greater stress, anxiety, depression, and burnout. This is harmonious to a survey conducted by the U.S. Department of Health and Human Services, where responses showed that 84% of respondents said their workplace conditions had contributed to at least one mental health challenge (H.H.S., 2024). These claims are further strengthened by Sleek (2024), who suggests that workplace toxicity can lead to several health issues, such as depression, heart disease, and cancer. The correlation between toxic workplace culture and mental health is formidable; moreover, Boddy (2014) takes a step further to claim that toxic employers cause the toxic environment, ultimately decreasing the wellbeing of others.

In addition to mental health, a toxic workplace culture also contributes to interpersonal conflict among employees, particularly between supervisors and subordinates. Interpersonal conflicts can manifest in various forms, including value conflicts, ego conflicts, and policy

conflicts (De Dreu & Gelfand, 2008). These conflicts often arise from differences in beliefs, communication styles, and expectations, creating a challenging environment for employees (Sonnentag, 2018). Such interpersonal conflicts have the greatest effect on subordinates, creating a direct correlation to a decrease in employee wellbeing (Chamberlain and Hodson, 2010). Subordinates, often feeling the weight of power dynamics, may experience heightened stress and anxiety, leading to decreased productivity and morale (Sonnentag, 2018).

Furthering the point on interpersonal conflict, cultural conflict within the workplace can exacerbate the challenges faced by both supervisors and employees, particularly when there are differences in cultural backgrounds and work expectations (Wang et al., 2004; Shao et al., 2013). For instance, a supervisor from a work-heavy cultural background may have become accustomed to a work culture mindset that values long working hours, rapid task completion, and hierarchical respect. This could clash with employees from cultures that prioritise work-life balance, deliberate decision-making, or more egalitarian workplace structures. Such cultural differences can lead to misunderstandings and conflicts in areas like pace, overtime expectations, and communication styles. For example, the supervisor may perceive an employee who works fewer hours or at a slower pace as lazy or uncommitted, while the employee might view the manager's expectations as unreasonable or exploitative. These cultural misalignments can create tension, reduce job satisfaction, and ultimately impact productivity (De Dreu & Gelfand, 2008). Addressing these cultural conflicts requires mutual understanding, communication, and compromise. Organisations need to foster an environment where diverse work styles and cultural perspectives are respected and valued, to create an environment that is culturally-inclusive and has less interpersonal and cultural conflict (Sonnentag, 2018).

Customer interaction can be another source of damage to wellbeing, in the form of emotional labour. This is because employees need to display emotions that are organisationally desirable in their interactions with customers (Chamberlain & Hodson, 2010). As these emotions are not always genuine, it can be emotionally taxing on employees to constantly keep up a facade. This ties in to mental health and wellbeing, as employees end up becoming separated from their genuine feelings. Thomas and Chaleff (2017) suggest that employees tend to follow a separate set of ethical standards at work than they do at home, emphasising how employees in the workforce tend to change themselves while working, often leading to a disconnect that can cause interpersonal conflict and harm mental health. Furthermore, effective conflict management strategies, such as improving communication and fostering supportive leadership, are essential to mitigate these issues and promote a healthier work environment (De Dreu & Gelfand, 2008). The effects that toxic workplace culture has extend beyond feelings of disdain towards their job but impact mental health in individual employees.

Influence of Toxic Culture on Employee Productivity

Relating the effects of toxic workplace culture to employee productivity, Amna Anjum et al. (2018) highlight that toxic workplace culture significantly undermines employee productivity, leading to a decline in overall job performance. Boddy (2014) concurs, noting correlations

between workplace bullying and counterproductive work behaviour. Bullying is a key factor that contributes to workplace toxicity, with counterproductive behaviour not only reducing productivity but actively harming the organisation. Anjum et al. (2018) further argue that, in addition to a direct negative correlation between toxic workplace culture and job productivity, there is an indirect relationship with job burnout, which ultimately decreases job satisfaction. Similarly, Bhatt et al. (2024) assert that a lack of necessary support, recognition, and a positive work environment inevitably leads to reduced productivity through the neglect of employee well-being.

While the general consensus is that toxic workplace culture creates significant negative impacts on employees, Gross, an editor at *Monster*, attempts a counter-argument by providing potential solutions to survive in a toxic workplace. Gross (2024) asserts that it is important to “*be nice to everyone, even to toxic co-workers.*” While this may seem plausible in theory, Chamberlain and Hodson (2010) caution that interpersonal conflict can stem from the constant disconnect of genuine emotions, ultimately leading to a decrease in mental health. Thus, by acting nice to toxic co-workers, individuals will find themselves disconnecting from their genuine emotions, lowering their own mental health. As a result, Gross’ argument is weakened as it fails to consider the impacts on non-toxic individuals who attempt to falsely amend relations with those who are toxic.

Solutions to Toxic Workplace Culture

Addressing toxic workplace culture requires a multifaceted approach centred on leadership and organisational change. Since toxic leadership is identified as the primary driving force of toxic culture, organisations should prioritise leadership training and development to foster positive management styles (Sull & Sull, 2022). This includes educating leaders on effective communication, trust-building, and conflict resolution techniques to prevent micromanagement and abusive behaviours (Gandolfi & Stone, 2022). Implementing open communication policies and feedback mechanisms can help create a safe environment where employees can express concerns without fear (Peetz, 2019). Recent studies demonstrate the effectiveness of targeted workplace programs in addressing toxic culture and improving employee well-being. By implementing training programs, focus groups and expert-led workshops to improve workplace culture, the result was lower employee turnover, denser support networks, and increased workplace satisfaction. Both studies highlight the importance of multi-faceted, evidence-based approaches in creating healthier work environments and improving employee productivity (Alan et al., 2021; Ammendolia et al., 2016).

Additionally, organisations should focus on improving employee well-being by establishing support systems such as employee assistance programs, promoting work-life balance, and addressing mental health concerns (H.H.S., 2024; Sleek, 2024). To further reinforce a healthy workplace culture, positive behaviours should be recognised and rewarded, and clear policies should be enacted against toxic actions (Baquero, 2023). Furthermore, fostering inclusivity and diversity can help reduce conflicts arising from differences in values and

communication styles (De Dreu & Gelfand, 2008). By implementing these solutions, organisations can create a healthier work environment that not only improves employee morale and well-being but also boosts overall productivity (Rasool et al., 2021). Studies conducted recently have shown that support systems such as employee assistance programs or programs that address mental health have decreased absenteeism and presenteeism. By focusing on employee well-being, organisations have found increases in productivity and satisfaction, ultimately leading to a healthier workplace culture (Richmond et al., 2017; Wagner et al., 2016).

Conclusion

In summary, corporate employee productivity is hampered by toxic workplace culture. As evidenced by numerous studies, the detrimental effects of toxic leadership, neglect of employee well-being, and interpersonal conflicts undermine individual performance, and indirectly affect organisations. The correlation between a toxic environment and increased employee turnover, mental health issues, and decreased job satisfaction calls for a need for organisations to find solutions to these challenges. Possible solutions that can foster a healthier workplace culture include effective leadership training, support systems for mental health and work-life balance, and recognizing and rewarding positive behaviours. This could look like comprehensive leadership development programs focused on emotional intelligence, trust and open communication or initiatives that promote employee well-being through mental health support and work-life balance. Additionally, by embracing inclusivity and diversity, organisations can mitigate conflicts arising from cultural differences, ultimately leading to enhanced employee morale and productivity. Re-evaluating performance management systems to include cultural contributions can help identify toxic behaviours early on and stop them from festering into a toxic culture. Future research should focus on long-term impact studies to assess the effects of toxic environments on employee health and performance, evaluate the effectiveness of various intervention strategies, and explore how emerging technologies can help detect and mitigate workplace toxicity. As organisations recognize the significant costs associated with toxic environments—both in terms of financial losses and employee well-being—they must act to cultivate a culture that values respect, collaboration, and support. Only through such efforts can these organisations hope to thrive in today's competitive landscape while ensuring the well-being of their most valuable asset: their employees.

Works Cited

- Alan, S., Corekcioglu, G., & Sutter, M. (2021). Improving Workplace Climate in Large Corporations: A Clustered Randomized Intervention. IZA Institute of Labor Economics Discussion Paper Series. <https://doi.org/10.2139/ssrn.3923565>.
- Ammendolia, C., Côté, P., Cancelliere, C., Cassidy, J., Hartvigsen, J., Boyle, E., Soklaridis, S., Stern, P., & Amick, B. (2016). Healthy and productive workers: using intervention mapping to design a workplace health promotion and wellness program to improve presenteeism. *BMC Public Health*, 16. <https://doi.org/10.1186/s12889-016-3843-x>.
- Anjum, A., Ming, X., Siddiqi, A. F., & Rasool, S. F. (2018). An Empirical Study Analyzing Job Productivity in Toxic Workplace Environments. *International Journal of Environmental Research and Public Health*, 15(5), 1035. <https://doi.org/10.3390/ijerph15051035>.
- Baquero A. (2023). Authentic Leadership, Employee Work Engagement, Trust in the Leader, and Workplace Well-Being: A Moderated Mediation Model. *Psychology research and behavior management*, 16, 1403–1424. <https://doi.org/10.2147/PRBM.S407672>
- Bhatt, J., Bordeaux, C., & Fisher, J. (2024). The workforce well-being imperative. Deloitte Insights. <https://www2.deloitte.com/us/en/insights/topics/talent/employee-wellbeing.html>
- Boddy, C. R. (2014). Corporate Psychopaths, Conflict, Employee Affective Well-Being and Counterproductive Work Behaviour. *Journal of Business Ethics*, 121(1), 107-121. www.jstor.org/stable/42921368.
- Chamberlain, L. J., & Hodson, R. (2010). "Toxic Work Environments: What Helps and What Hurts." *Sociological Perspectives*, 53(4), 455-478. <https://doi.org/10.1525/sop.2010.53.4.455>.
- De Dreu, C. K. W., & Gelfand, M. J. (Eds.). (2008). The psychology of conflict and conflict management in organizations. Psychology Press. <https://doi.org/10.4324/9780203810125>
- Gandolfi, F., & Stone, S. (2022). Toxic Leadership: Behaviors, Characteristics, and Consequences. *California Institute of Advanced Management*. <https://www.ciam.edu/toxic-leadership-behaviors-characteristics-and-consequences>
- Gross, E. L. (2024). How to Survive a Toxic Workplace. *Monster*, www.monster.com/career-advice/article/handle-toxic-workplace.
- Healy, S., & Dowell, R. S. (2017). Cleaning up Toxic Workplaces. *GPSolo*, 34(4), 28-31. www.jstor.org/stable/26425919.
- Looman, N., van Woezik, T., van Asselt, D., Scherpbier-de Haan, N., Fluit, C., & de Graaf, J. (2022). Exploring power dynamics and their impact on intraprofessional learning. *Medical education*, 56(4), 444–455. <https://doi.org/10.1111/medu.14706>
- McCulloch, A. (2016). Toxic work environments [Doctoral dissertation, Carleton University]. Carleton University Research Repository. <https://repository.library.carleton.ca/downloads/nc580n548>
- Peetz, D. (2019). Management, culture and control. In *The Realities and Futures of Work* (pp. 113–140). ANU Press. <http://www.jstor.org/stable/j.ctvq4c16w.10>
- Rasool, S. F., Wang, M., Tang, M., Saeed, A., & Iqbal, J. (2021). How Toxic Workplace Environment Effects the Employee Engagement: The Mediating Role of Organizational

- Support and Employee Wellbeing. *International Journal of Environmental Research and Public Health*, 18(5), 2294. <https://doi.org/10.3390/ijerph18052294>
- Richmond, M., Pampel, F., Wood, R., & Nunes, A. (2017). The Impact of Employee Assistance Services on Workplace Outcomes: Results of a Prospective, Quasi-Experimental Study. *Journal of Occupational Health Psychology*, 22, 170–179. <https://doi.org/10.1037/ocp0000018>.
- Shao, R., Rupp, D., Skarlicki, D., & Jones, K. (2013). Employee Justice Across Cultures. *Journal of Management*, 39, 263 - 301. <https://doi.org/10.1177/0149206311422447>.
- Singh, S., Solkhe, A., & Gautam, P. (2022). What do we know about Employee Productivity?: Insights from Bibliometric Analysis. *Journal of Scientometric Research*, 11(2), 183-198. <https://doi.org/10.5530/jscires.11.2.20>
- Sleek, S. (2024). Toxic workplaces leave employees sick, scared, and looking for an exit. How to combat unhealthy conditions. *American Psychological Association*. <https://www.apa.org/topics/healthy-workplaces/toxic-workplace>
- Sonnentag, S. (2018). The recovery paradox: Portraying the complex interplay between job stressors, lack of recovery, and poor well-being. *Research in Organizational Behavior*, 38, 169-185. <https://doi.org/10.1016/j.riob.2018.11.002>
- Stoltz, R. (2023). Toxic Work Culture: Signs of an Unhealthy Work Environment & How to Fix It. *workhuman*. www.workhuman.com/blog/toxic-work-culture-environment/.
- Sull, D., & Sull, C. (2022). How to Fix a Toxic Culture. *MIT Sloan Management Review*. sloanreview.mit.edu/article/how-to-fix-a-toxic-culture/.
- Sull, D., Sull, C., & Zweig, B. (2022). Toxic Culture Is Driving the Great Resignation. *MIT Sloan Management Review*. sloanreview.mit.edu/article/toxic-culture-is-driving-the-great-resignation/.
- Thomas, Ted, and Ira Chaleff. (2017). Moral Courage and Intelligent Disobedience. *InterAgency Journal*, 8(1), 58-66. <https://thesimonscenter.org/wp-content/uploads/2017/03/IAJ-8-1-Winter2017-pg58-66.pdf>.
- U.S. Department of Health and Human Services. (2024) Workplace Well-Being. *U.S. Department of Health and Human Services*. www.hhs.gov/surgeongeneral/priorities/workplace-well-being/index.html#framework.
- Wagner, S., Koehn, C., White, M., Harder, H., Schultz, I., Williams-Whitt, K., Wärje, O., Dionne, C., Koehoorn, M., Pasca, R., Hsu, V., McGuire, L., Schulz, W., Kube, D., & Wright, M. (2016). Mental Health Interventions in the Workplace and Work Outcomes: A Best-Evidence Synthesis of Systematic Reviews. *The International Journal of Occupational and Environmental Medicine*, 7, 1 - 14. <https://doi.org/10.15171/ijoem.2016.607>.
- Wang, P., Lawler, J., Walumbwa, F., & Shi, K. (2004). Work-family conflict and job withdrawal intentions: The moderating effect of cultural differences. *International Journal of Stress Management*, 11, 392-412. <https://doi.org/10.1037/1072-5245.11.4.392>.
- Wells, V. (2022). Workers are stressing out over their finances — and it's costing employers billions. *Financial Post*. financialpost.com/fp-work/worker-financial-stress-employers-cost-billions.

Examining the Reading Brain By Jasmine Ndula

Abstract

Reading is one of the most ancient methods of acquiring knowledge, dating back to 3,000 BCE; it involves decoding letters and symbols by the brain, further translated to speech and writing. Moreover, reading is a crucial aspect of learning, requiring refined cognition and phonological skills to grasp and further apply the attained knowledge. Research in cognitive science has played a large role in shaping our current understanding of the reading process. The history of studying how reading processes function started with Wilhelm Wundt, as part of cognitive psychology. Psychologists studying the cognition behind reading comprehension suggest that reading involves creating a mental image of the text, which forms a memory representation of the text information for interpretation. It is no surprise, then, that there is a strong empirical relationship between reading and memory. This review aims to understand the science behind brain functions while reading and its implications on learning and memory.

Cognitive Processes of Reading

To read, one must “decode” the displayed text or symbols. The decoding process entails the “ mapping of visual (orthographic) information onto auditory (phonological) and conceptual (semantic) representations” (Turkeltaub et al.1) by the brain. This establishes that the reading process is procedural and takes different skill levels to be carried out efficiently; it starts with phonemic awareness, the visual recognition of letters (Kweldju) whereby a reader learns to correlate spelling to sound; this forms the basis of the decoding process. As skill level increases through exposure to various words, word recognition becomes an automatic process that is built upon and is highly dependent on overall language skill level, exposure to vocabulary, and familiarity with syntactic structures (Cunningham and Stanovich).

Cunningham and Stanovich argue that the amount of reading an individual accomplishes has significant implications on the ease of comprehension. They argue that “Children's books have 50 percent more rare words in them than does adult prime-time television and the conversation of college graduates” (3). Thus, although television and other sources of vocabulary have the capacity for new word learning, it is not as significant as the capacity presented by reading books; reading presents a more efficient pathway for building comprehension skills (Cunningham and Stanovich). Reading books is also shown to have an impact on student reading fluency levels (DeWaard). Reading fluency involves the automaticity of word recognition, prosody (the appropriate intonation and pitch), and accuracy (the correct pronunciation) for effective comprehension (DeWaard). By exposing readers to various books and reading-related activities, fluency in reading is enhanced significantly (DeWaard).

Areas of the Brain Associated with Reading

Reading is more than an instant text-to-sound-to-meaning; it takes brain faculties to piece information together and process it to produce meaningful outcomes such as speech and writing

(Landi et al.). The brain utilizes four distinct regions associated with reading: the frontal, parietal, temporal, and occipital lobes which work together to form the entire reading process (Kearns et al.). To investigate brain activity during the reading process, researchers use fMRI (functional magnetic resonance imaging) a brain imaging technique that measures the amount of oxygen supplied to active parts of the brain; the more oxygen provided to a specific area, the more brain activity (Kearns et al.).

The brain's left hemisphere is largely associated with analytical skills such as language processing and visual representation (Teo et al.); it comprises the temporoparietal system at the dorsal end, the occipitotemporal region at the ventral end, and the inferior frontal gyrus at its anterior (Landi et al.). See Figure 1 below.

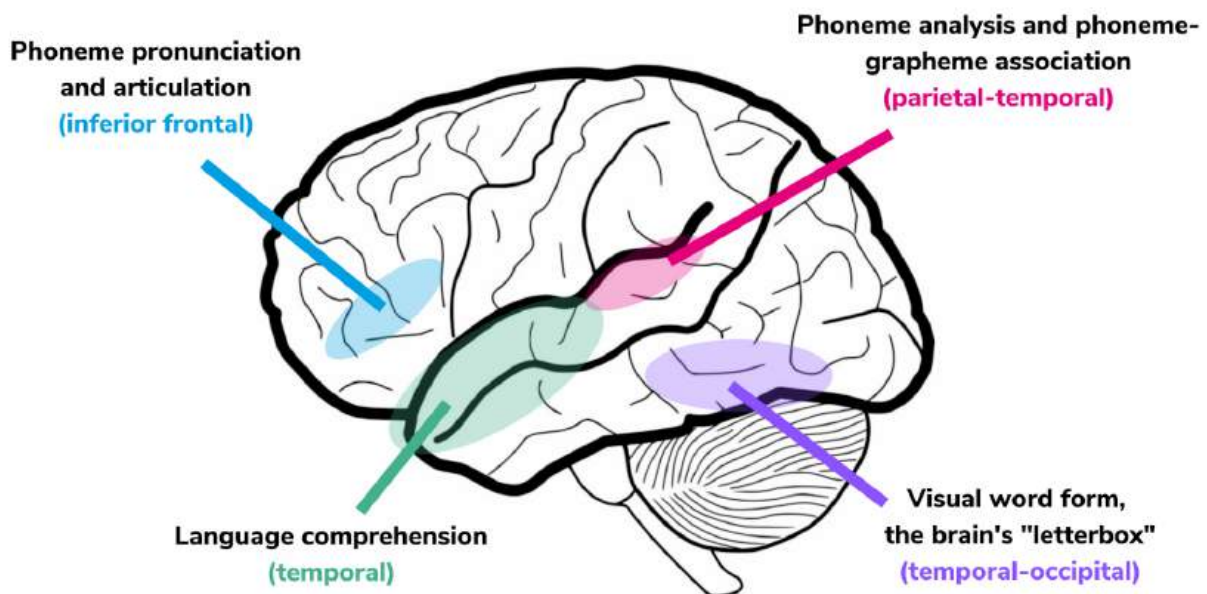


Figure 1. *Areas of the brain associated with reading.*
<https://relay.libguides.com/science-of-teaching-reading-resource-guide>

The inferior frontal gyrus is an essential region for reading, as it forms a basis for word recognition (Landi et al.). What this means is that a word can be automatically recognized by decoding its phonetics. Cornelissen et al. find that the left inferior frontal gyrus region is activated within the first 200 milliseconds (ms) of reading; these findings are essential in understanding the role of the inferior frontal gyrus in linking sounds to their text representation during the initial stages of reading. Kearns suggests that the inferior frontal gyrus could be essential in the sequencing of information such that it “may help readers to put sounds in the correct order when they are ready to say it aloud.”(8)

The temporoparietal region is stimulated for language comprehension (Kweldju). This region comprises three essential areas of interest to reading researchers; the supramarginal gyrus, the superior temporal gyrus, and the angular gyrus. The supramarginal gyrus has been linked to language perception and processing while the superior temporal gyrus works as a speech processor (Kweldju; Kearns). The angular gyrus is stimulated to process semantic representations of words; these regions work together to perceive language and syntax structures and derive meaning from them (Segheir et al.)

The ventral occipitotemporal gyrus (vOT) “is involved in the perception of visually presented objects and written words” (Price and Devlin 1). In the early stages of learning to read, there is increased activity in this region as the brain gives priority to potentially meaningful information (Price and Delvin); the occipitotemporal gyrus consists of the fusiform gyrus at the lateral end that implements high-level visual processing which includes face perception, object recognition as well as reading (Weiner and Zilles). The occipitotemporal gyrus is the brain’s word processor as it retains previously learned letters and symbols and recognizes them when stimulated (Price and Delvin). Together, these regions collaborate to create the reading flow (Coltheart).

Impact of Reading on the Brain

Ample research shows that reading has a positive effect on the brain for memory and strengthening neural networks (Ferreira et al.). To be more precise, brain faculties become stronger as a result of engaging in reading activities simply because such activity demands for processing of information to create new neural pathways for newly learned material. This is extremely beneficial for learners and can potentially divert neurodegenerative diseases down the road (Stine-Morrow).

Strengthening Neural Networks

Reading and reading-related tasks have the potential to reorganize brain networks (Houde et al.). For example, Wang et al. demonstrate that learning to read Chinese results in the development of thalamo-occipital and fronto-striatal circuits. The thalamo-occipital circuit is essential in creating a visual pathway for acquired information while the frontal-striatal circuit is important for memory and attention during reading (Wang et al.). Further, Wang et al. suggest that the thalamo-occipital pathway is strengthened during the initial stages of learning to read but slowly deteriorates when reading ability becomes better. In contrast, the fronto-striatal pathway increases in activity as reading proficiency develops; this is because the ease of visual recognition of words increases with reading practice and the thalamo-occipital pathway is not activated for learning a new skill. Memory and attention are also heightened with reading practice, so the frontal-striatal pathway is activated more frequently. To further demonstrate the important ways that reading creates new neural pathways, Skeide finds that a six-month literacy training for people who cannot read, increased neuroplastic connections “between the occipital lobe (visual word recognition area) and subcortical areas in the midbrain and thalamus” (1). By

the formation and strengthening of neuroplastic pathways, reading became an easy task for the initially illiterate participants enabling them to read in only six months (Skeide).

Enhancing Memory

Reading and memory go hand in hand. To learn something from a text one must remember it first and then correlate it to other parts of the text to derive a wholesome meaning (Elfenbein); this is the function of the working memory. Working memory is the “retention of a small amount of information in a readily accessible form that facilitates planning, comprehension, reasoning, and problem-solving.” (Cowan 1). Cowan’s definition explains that working memory is a processing unit tasked to retrieve and utilize necessary information for a meaningful outcome. This definition aligns with several others as many of them explain that working memory has limited capacity. However, building literacy habits can potentially expand its capacity. Ferreira et al. suggest that with increasing literacy levels comes a corresponding increase in the capacity of working memory. According to their study, “students in the 3rd and 4th year showed better results in the physical assessment of working memory as well as in the evaluation of reading level” (1) when compared to students in earlier grades of elementary school. Notably, the authors also point out that one of the study’s limitations is that reading levels and working memory may not directly correspond to each other due to other factors at play during the development of reading (Ferreira et al.).

Literacy practices can enhance memory by strengthening regions in the brain involved with memory (Stine-Morrow et al.). For instance, sustained literacy engagement has been known to “buffer the clinical manifestations of Alzheimer’s disease” (Stine-Morrow et al. 3) by strengthening neural pathways associated with reading and memory. Sustained literacy engagement means that reading is conducted continuously throughout one’s lifetime even after education. Continuous exposure to literature enhances the brain’s neuroplasticity by extensively engaging working memory, potentially staggering the risk of such neurodegenerative diseases (Stine-Morrow et al.)

Works Cited

- Coltheart, Max. "Dual route and connectionist models of reading: An overview." *London review of education* 4.1 (2006).
- Cowan, Nelson. "Working memory underpins cognitive development, learning, and education." *Educational psychology review* 26 (2014): 197-223.
- Cornelissen, Piers L., et al. "Activation of the left inferior frontal gyrus in the first 200 ms of reading: evidence from magnetoencephalography (MEG)." *PloS one* 4.4 (2009): e5359.
- Cunningham, Anne E., and Keith E. Stanovich. "What reading does for the mind." *Abstracts of scientific articles* 4389 (1998): 128-0.
- DeWaard, Jessica. "Effective reading strategies to increase reading fluency." (2021).
- Elfenbein, Andrew. "Cognitive science and the history of reading." *Pmla* 121.2 (2006): 484-502.
- Ferreira, T. de L., C. M. T. Valentin, and S. M. Ciasca. "Working memory and reading development." *Psychology* 4.10 (2013): 7.
- Houdé, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: an fMRI meta-analysis of 52 studies including 842 children. *Developmental science*, 13(6), 876-885.
- Kearns, Devin M., et al. "The neurobiology of dyslexia." *Teaching Exceptional Children* 51.3 (2019): 175-188.
- Kweldju, Siusana. "Neurobiology Research Findings: How the Brain Works during Reading." *Pasaa: Journal of language teaching and learning in Thailand* 50 (2015): 125-142.
- Landi, Nicole, et al. "Neurobiological bases of reading comprehension: Insights from neuroimaging studies of word-level and text-level processing in skilled and impaired readers." *Reading & Writing Quarterly* 29.2 (2013): 145-167.
- Price, Cathy J., and Joseph T. Devlin. "The interactive account of ventral occipitotemporal contributions to reading." *Trends in cognitive sciences* 15.6 (2011): 246-253.
- Skeide, M. A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., ... & Huettig, F. (2017). Learning to read alters cortico-subcortical cross-talk in the visual system of illiterates. *Science advances*, 3(5), e1602612.
- Stine-Morrow, Elizabeth AL, et al. "The effects of sustained literacy engagement on cognition and sentence processing among older adults." *Frontiers in Psychology* 13 (2022): 923795.
- Seghier, Mohamed L., Elizabeth Fagan, and Cathy J. Price. "Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network." *Journal of Neuroscience* 30.50 (2010): 16809-16817.
- Teo, Wei-Peng, David White, and Helen Macpherson. "Using noninvasive methods to drive brain-computer interface (BCI): the role of electroencephalography and functional near-infrared spectroscopy in BCI." *Smart Wheelchairs and Brain-Computer Interfaces*. Academic Press, 2018. 33-63.
- Turkeltaub, Peter E., et al. "The neural basis of hyperlexic reading: An fMRI case study." *Neuron* 41.1 (2004): 11-25.

- Weiner, K. S., & Zilles, K. (2016). The anatomical and functional specialization of the fusiform gyrus. *Neuropsychologia*, 83, 48–62.
- Wang, Y., Luo, J., Ma, L., Chen, R., Wang, J., Chu, C., ... & Tao, S. (2022). Learning to read Chinese promotes two cortico-subcortical pathways: The development of thalamo-occipital and fronto-striatal circuits. *Frontiers in Neuroscience*, 16, 983084

Politics and Paternalism: The Conditions Behind Capabilities By Matthew Jeong

Abstract

The concept of political justice has been defined loosely as the virtue of a society, or the virtue by which the state provides for citizens. This essay turns to a particular framework of political justice, the capabilities view. Theories on the capabilities view have been debated by different philosophers such as Martha Nussbaum and Rosa Terlazzo. The primary points of contention are on whether the conditions for the capabilities view actually providing for individuals override freedom to the point of tyranny. Nussbaum argues that her version of the capabilities view provides full capabilities for everyone with limited government intervention, while Terlazzo disagrees and states that it is only possible to commit to one or the other. This essay proceeds to create a new notation and requirements for the capabilities view that can provide an alternative response to Nussbaum and Terlazzo's critique. The aim is to describe the ultimate goal of political justice as allowing individuals to achieve capabilities, and how this framework can be built while allowing for freedom of choice.

1. Introduction

The nature of political justice mainly concerns the success of the governing body, especially in its treatment of citizens. Out of those, one framework known as the capabilities view argues that having the freedom to choose whatever action or state is the end goal of political justice.

One of the questions confronting capabilities theorists is what it takes for individuals to have a given capability. On one view, defended by Rosa Terlazzo, to have a capability requires breath comprehensiveness along with a rejection of political liberalism. Political liberalism in this instance connotes a measure of political neutrality, in which political neutrality and a lack of government intervention in individual life is observed. However, Terlazzo rejects this claim and argues that in order for the capabilities view to actually provide capabilities to individuals, it necessitates distancing from political neutrality. She argues that the conditions described by Nussbaum mean that internal capabilities, or people's own desires, can withhold them from achieving their full capabilities which would require government intervention.¹ But in this paper, I make alterations to Terlazzo's critique, arguing instead that capabilities require deliberative options backed up by several conditions. One might then wonder whether the adoption of such a constraint on capabilities would nevertheless lead to problems associated with paternalism. In this paper, I will further argue that, while some paternalism is necessary in accordance with the capabilities view, such paternalistic interventions need not rise to a level that would pose a serious problem for people's freedoms.

In section II, I'll begin by laying out some of the important features of the capabilities view, before turning to Terlazzo's critique, aimed at showing that this view can't be paired with

¹ Rosa Terlazzo, *How Politically Liberal Should the Capability Approach Want to Be?* (Journal of Political Philosophy 25, no. 3, 2017)

political liberalism. Then, in section III, I'll turn to my own argument, showing that Terlazzo might be mistaken in her reasons for thinking that there's this incompatibility, but she's not completely wrong in her assessment. The capabilities approach needs paternalism. My argument will hinge on claims on deliberative options and capabilities. But I'll also argue that this needn't be a problem for the capabilities view, before turning to two key objections to my argument. The first objection is that my argument doesn't fulfill the sufficient conditions whereas the second takes aim at the necessary conditions. Ultimately, I will show that these objections are premised on the problematic assumptions that hinge on the conditions being unfulfilled respectively, and ought to be rejected.

2. The Capabilities View and Terlazzo's Critique

2.1 The Capabilities View

According to the capabilities framework, developed by philosophers such as Martha Nussbaum, the success of society's governing body is measured by the access citizens have to capabilities, with certain central capabilities being vital. The state's role is to ensure that there are no obstacles preventing citizens from having those capabilities.²

While different capabilities theorists might have different views about what capabilities are relevant to political justice, Nussbaum provides a list of ten central capabilities.³ These include:

1. LIFE. Being able to live a natural human lifespan without premature death or having the quality of life reduced significantly to the point of not worth living.
2. BODILY HEALTH. In good physical and mental health, nourishment, and housing
3. BODILY INTEGRITY.
4. SENSES, IMAGINATION, AND THOUGHT.
5. EMOTIONS.
6. PRACTICAL REASON. Being able to reason and think critically and decide on particular viewpoints according to one's free will. This includes being able to freely rationalize and requires protection of religious observance as well.
7. AFFILIATION.
 - A. FRIENDSHIP.
 - B. RESPECT.
8. OTHER SPECIES.
9. PLAY.
10. CONTROL OVER ONE'S ENVIRONMENT.
 - A. POLITICAL.
 - B. MATERIAL.

² For one place where Nussbaum develops this view, see her "Capabilities and Human Rights."

³ For Nussbaum's discussion of these capabilities, see "Capabilities and Human Rights," pp. 287-288

This contrasts with other major views of political justice, like the equal distribution of resources as ascribed by Marxism in which class is nonexistent, or Dworkin's political philosophy of liberalism in which the government must provide equal concern for each citizen. While these views have mainly to do with economic and social factors such as classism, the capabilities view also isn't measured by terms such as GDP or other common indicators. Instead, it's unique in the fact that the outcome of a societal measure or change is not the end goal, but the opportunities that people possess.

Nussbaum believes that people having the capability to pursue certain actions or being in certain states determines whether or not political justice is properly meted out in society. She has a list of ten central capabilities, such as bodily integrity (which includes being able to go where one wants, being safe from violent assault, etc. Nussbaum believes that this theory can be committed alongside political liberalism, which holds values in maintaining neutrality and is against governmental overreach.

Capabilities theorists distinguish between the possession of a capability, on the one hand, and functioning, or the exercise of the capability, on the other. This is often because capabilities theorists are concerned with maintaining some role for freedom of choice, or the freedom to choose whether or not to exercise a capability. A capability is understood as the opportunity to be in a certain state, or to engage in certain activities. Capabilities cover a broad spectrum of components- basic capabilities can include actions like having the freedom of movement, while specific realizations of that capability could be resources like access to vehicles. A functioning, on the other hand, is having achieved an action or activity. It's a debate in the literature, to which I'll return below, the extent to which capabilities theorists are also committed to aiming at certain functionings, too. For example, it's been suggested that certain capabilities depend upon realizing certain achievements.⁴ Afschin Gandjour, for instance, argues that certain functionings must be met for certain capabilities to exist in the first place, such as in the scenario of mental health requiring prerequisites like happiness.⁵ Ingrid Robeyns brings up the idea of welfare states, in which people are given a base level of economic stability in order to pursue other capabilities such as having housing or seeking employment.⁶

Finally, another crucial aspect of Nussbaum's thinking about her capabilities view is that it is compatible with political liberalism. Political liberalism in this instance equates to political neutrality, in which paternalism stays as minimized as possible and avoids interfering in individuals' lives. She argues that public policy should be aimed at allowing everyone to achieve basic internal capabilities, states in which sufficient conditions are met for one to exercise that capability. However, actually exercising those capabilities would be unaided, allowing for political liberalism in this sense. Nussbaum argues the following about the capabilities view, "As

⁴ For further discussion of these capabilities, see [Afschin Gandjour,] "Mutual dependency between capabilities and functionings in Amartya Sen's capability approach," pp. 347.

⁵ For further discussion of these capabilities, see [Afschin Gandjour,] "Mutual dependency between capabilities and functionings in Amartya Sen's capability approach," pp. 347.

⁶ For further discussion of these capabilities, see [Ingrid Robeyns,] "Functionings or Capabilities?," pp. 341.

for the argument from paternalism, it nudges us strongly in the direction of what might be called political rather than comprehensive liberalism, in the sense that it urges us to respect the many different conceptions of the good citizens may have and to foster a political climate in which they will each be able to pursue the good (whether religious or ethical) according to their own lights, so long as they do no harm to others.”⁷ Essentially, political liberalism in Nussbaum’s capabilities view has to do with liberty and self-determination, the power of self-responsibility over paternalistic intervention.

2.2 Terlazzo’s Critique

However, there are those who criticize Nussbaum’s approach. For example, Terlazzo argues that political liberalism is not compatible with the capabilities view, for Terlazzo believes that certain internal obstacles – not addressed by Nussbaum – prevent people from accessing their full capabilities.

Terlazzo is specifically concerned about certain examples of agents who would not appear to have certain of the central capabilities Nussbaum invokes, due to what she describes as particular “intrinsic” obstacles that Nussbaum does not discuss. She goes on to speak on cases where the obstacles come into play: “And cases like this suggest that we should also recognize nearby cases, in which a person feels drawn to exercise some capability, but cannot bring himself to do so because he doesn’t believe that it is appropriate for someone like him to behave in that way. Think here of the self-described manly man who yearns to play jazz saxophone, but thinks that doing so is effeminate. Even if the norms are not strictly externally enforced, he may have internalized them in a way that renders the option off limits as far as he is concerned.” She goes on to argue as follows:⁸

- A. A.1. According to proponents of the capability view, political justice is measured in terms of people’s actual capabilities – that is, their opportunities for being or doing certain things.
- A. A.1. Internal obstacles can also prevent people from accessing their capabilities.
- A. A.1. The capability view must decide on what people need to value in order for them to have access to those capabilities that may go against their values.
- A. A.1. Incorporating requirements on such values would go against political liberalism’s rejection of breath-wise broad comprehensive political doctrine.

Therefore, the capability is incompatible with political liberalism. (From Premises 3, 4)
As in Premise 1, political justice is measured strictly by capabilities. However, as mentioned in Premise 2, there are roadblocks that can limit people from having those capabilities. For

⁷ Nussbaum, “Women and Human Development,” pp. 50.

⁸ Terlazzo, “How Politically Liberal Should the Capabilities Approach Want to Be?” pp. 292-293.

example, a Haredi Jewish woman would likely not become an aerospace engineer due to her values. Since becoming an engineer would never occur to her as an option in the first place, she would not have the capability to become an engineer. However, as mentioned in Premise 3, people have different values, some may have values that prevent them from behaving or thinking a certain way. For the woman to have the capability to become an engineer, they would have to go against political liberalism (mentioned in premise 4) and decide their values for them, or the capabilities view would need to be reduced to merely having functionings, or achievements as its goal. In other words, Terlazzo states that to get rid of internal obstacles such as those mentioned, political liberalism would have to be taken out of consideration, meaning the capabilities view and political liberalism are not compatible.⁹

3. Reply to Terlazzo

Terlazzo's critique, when analyzed further, does seem to run into an issue. Though I agree partially with Terlazzo in that the capabilities view is not completely compatible with practical applications of political liberalism, I take issue with the notion of capability implicit in Terlazzo's case, as it is much too stringent. When analyzed, there seems to be too general of a reach, which I point out below:

1. Suppose that Terlazzo is right that internal obstacles are created as a consequence of people's actual motivations and thoughts. (Assumption)
1. But, people's actual motivations and thoughts are dynamic and can change easily, which means that there is no set way to determine what people's actual deliberative options might be. (Assumption: fact based on psychology)
1. If we accept Terlazzo's view, then, we would be committed to a view of capability that would render whether someone had the capability basically always changeable and/or unknowable, which would make it useless. (From 1, 2)
1. Therefore, we should reject Terlazzo's account of what's required for full capability. (From 3)

Terlazzo's argument seems to reach too far in its critique of Nussbaum's view. Terlazzo claims that internal obstacles are a result of people's own thought processes. For example, an uncontacted tribal member with little concept of modern-day finances could never think of becoming a stock broker. However, this argument covers too many variables. Thoughts and opinions are fluid and easily changeable. A person might dislike something one day and love it the next day. Due to the dynamic nature of motivations and thoughts, there isn't a concrete way

⁹ Terlazzo, "How Politically Liberal Should the Capabilities Approach Want to Be?" pp. 300-312.

to narrow down the internal obstacles and options that are available to a person. For example, a boy may hate basketball, and then decide that he enjoys it the next day after a particularly fun game. The capability for playing basketball had changed as did the internal obstacles in a manner that is much too mercurial. Terlazzo's argument would mean that access to capabilities could be there one moment and gone the next, changing much too quickly and unpredictably to be of any practical use.

For example, in current psychology, one of the debates very much alive today is over whether people can be said to have stable character traits, including stable motivations. Situationists – who argue that people's situations, not traits, are the primary factor for their actions – point to certain studies that seem to confirm that people's actual motivations are highly dependent upon the situations in which they find themselves. For example, consider the Darley and Batson experiment, in which people were faced with competing demands for help, from a person slumped at a door needing assistance and from the experimenters who asked them to go to a location. Whether the experimenters asked them to hurry or not. Sixty-three percent of those with about five minutes to spare helped; 45 percent of those told to "go right over" and 10 percent of those who had been told they were already late assisted in helping the person in the doorway. The only factor that had changed was whether or not the experimenters told them to hurry- meaning that motivations and thought processes are often fluid, dependent on situations, and not at all completely internal or set.¹⁰

Such studies cast doubt on whether motivations can be considered a stable enough factor to change people's decisions and capabilities when other variables like situations can change them in an instant. And if motivations are not a stable character trait that can decide on a person's views and actions, then it remains impractical to consider as Terlazzo suggests in her critique of Nussbaum's capabilities view.

4. The Capabilities View and Paternalism

4.1 A New Capability Condition

We saw in the preceding section how Terlazzo argues for the incompatibility of the capabilities approach and political liberalism. Whereas I think Terlazzo has a point, nevertheless I take issue with her argument. In this section, I will do two things: first, show where Terlazzo's argument goes awry; second, provide an alternative argument to the same conclusion. I'll conclude this section by outlining those places where paternalism, then, would be warranted on the capabilities view, and argue that it shouldn't be allowed to extend beyond: capability should be prioritized over functioning when it's possible to do so.

Though Terlazzo's critique is spread too thin to be of realistic use, she does have a point in her introduction of internal obstacles. There are obstacles that prevent people from gaining access to capabilities that can be largely internal. The example of the Haredi woman as

¹⁰ This experiment and others are discussed in Kamtekar, "Situationism and Virtue Ethics on the Content of Our Character," pp. 462-464.

previously mentioned comes into play- it is true that she would not have the capability to become an aerospace engineer as she is now. This would require a revision of what it means to have access to a capability in the first place, which I show below. By using such a framework, an agent would have the capability for an action X only if it appears among their deliberative options. Furthermore, to appear among an agent's deliberative options, certain conditions could be met. Being exposed to agents who are similar to them committing actions X and being aware of taking X as an action would mean the agent would have the knowledge and awareness of X that causes them to develop X as an option. Outside circumstances and internal abilities must also match- but it should be in their point of view and belief.

This suggests the following two-part condition:

- (A) An agent A has the capability to X only if [necessary condition] X-ing could appear among A's deliberative options.
- (B) X-ing could appear among A's deliberative options if [sufficient condition]
 - 1) A is exposed to other agents who are (i) analogously like A, and who (ii) do X, AND
 - 2) A is aware of X-ing being an action they can take, AND
 - 3) A perceives that external circumstances are suitable to A's X-ing (e.g. there's nothing blocking A, AND A perceives that they have the external resources needed for X-ing, etc.), AND
 - 4) A perceives that they are internally equipped with the necessary intrinsic abilities, whatever they might be (e.g. bodily, psychological, etc.)

To understand this condition, let us consider a couple of examples. First, in the instance of women in Saudi Arabia prior to 2018, they did not have the capability to drive. Firstly, they did not have the capability to drive because it did not appear among their deliberative options. Women were not exposed to other women who drove, and many were not aware that driving could be an action they could take. External circumstances were also unsuitable: women were legally barred from being able to drive. They did however possess the intrinsic abilities to drive a car. However, as the other conditions were not met, driving could not appear among women's deliberative options and so they did not have the capability for it.

On the other hand, a healthy, athletic boy from an affluent family would have the capability to play professional soccer as it appears among his deliberative options. The boy is exposed to those like him who play soccer and is aware that he can play soccer. His family is well off so external circumstances are met, and he has the bodily and mental capacity to train and play soccer.

4.2 Objections and Replies

Now I will consider two objections to the above condition, with the first targeting the credibility of deliberative options and the second targeting a possible aversion to the capability of

an individual. I will then respond to the objections to argue that my conditions for the capabilities view still hold up.

First, one might object to the necessary condition using a case that purports to show an agent who has some capability, but for whom exercising that capability could not appear in their deliberative options. This might be seen to happen when agents' background influences their choices into predominantly choosing one option despite being aware of other opportunities available. For example, a sociopath has the capability for personal gain but does not have the deliberative options to do so as it is in their nature. In this case, wouldn't this present an issue where Condition A would be rendered as incorrect?

In response to this objection, I would argue that the sociopath in question does not actually have the capability for personal gain. The example of a sociopath who has the physical and mental ability for personal gain but does not consider it due to their nature actually highlights that true capability encompasses more than just physical or cognitive ability; it also includes the agent's deliberative framework. Capability is not just about the potential to perform an action, but about the integration of that potential within the agent's decision-making processes. Though the sociopath may technically be able to complete functionings of personal gain, he does truly possess the capability for it.

There is a second objection one might make, this time to the sufficient condition I have provided. One might argue that some agent could meet all of the conditions (1)-(4) with regard to the exercise of some capability, while it might still not be the case that the exercise of that capability could appear among that agent's deliberative options. If true, this would show that the sufficient condition I've attempted to give is not so sufficient after all. Consider as an example how some people are instinctively averse to certain actions, and so they could not have the deliberative options to perform those actions even (arguably) when the sufficient conditions I have given are met. Take this scenario: Xing=basketball (in a regular rec league) Agent A (absolutely hates losing, and would be a terrible basketball player, but technically can play). It might further be suggested that the agent in question meets all conditions (1)-(4). Does that mean that playing rec basketball could appear among Agent A's deliberative options? According to the condition, Agent A does. However, something is amiss: even if Agent A could in the off-chance consider playing, he is so averse to basketball that it's not really a real deliberative option, even when it does occur to Agent A to consider it.

In response to this objection, I would argue that this agent actually does not meet the sufficient condition, as specified. This is because the presence of aversion to the activity renders the agent such that they do not meet Condition 4. Being able to see one's self as being internally equipped to engage in some activity requires being able to see one's self engage in that activity in such a way that one is not in some kind of distress. It seems as though as long as the agent perceives themselves of being capable of being equipped with the facilities to play basketball, the agent does have the capability to do so. If the agent perceives themselves as not possessing the mental faculties to play basketball, then he would not have the capability to do so, which still fits in with the conditions I have listed. Motivation is yet again too fluid of an argument (as listed in

my response to Terlazzo) to remain a practical concern in a wide range of situations, but is still addressed somewhat in my version of capabilities as the perception of capability is what is important.

4.3 Resulting Paternalistic Commitments

If these conditions for these capabilities must be met in order for people to attain political justice, then paternalism is necessary for a society to function. The state would have a role to provide aid so that the conditions are met in a manner that doesn't overreach into power or interfere with freedom of choice- which is necessary for having the capability in the first place. A government's intervention to ensure that each of these four conditions is met might involve overlapping strategies, where - for example - work done to ensure that condition (3) is met will also help with the meeting of condition (1), etc.

To ensure that conditions (1) and (2) are met, governments would want to encourage the development of, or promote, certain ideas and ways of life involving people X-ing. As an example of this, governments could send representatives into schools to encourage and raise awareness for youths to use their power to vote.

To ensure that condition (3) is met, governments would want to create a conducive environment for the perception of external resources and circumstances favorable to X-ing, such that people would not be blocked from, but rather supported in, X-ing were they to choose to do so. For example, this might take the form of funding schools to provide educational guidance and resources assisting youths to find employment.

Finally, to ensure that condition (4) is met, governments would want to work to maintain or assist in the intrinsic abilities of people for X-ing such that people's perception of their intrinsic abilities for X-ing is sufficient. As an example, consider how government programs could be used to help fund minorities and women into STEM fields to convince and assist them that they have the intrinsic abilities to work in STEM.

4.4 Why interventions (A)-(B) are non-problematic

Having considered what government intervention to ensure capabilities would have to involve, I now turn to the question of whether the government's intervening along such lines would be a problem. Here, I'm especially interested in the question of whether it would interfere with people's freedom.

First, I do not think that there would be anything necessarily problematic with a government trying to intervene along lines (A)-(C). The first thing to say is that implementing these interventions does not necessarily violate free choice. None of these conditions specifically mention replacing certain options or ideas or applying favoritism, which would violate freedom of will. In fact, far from undermining free choice, such interventions can be seen as enhancing freedom, by giving people knowledge of what their options in life are.

Moreover, adopting such interventions does not necessarily require the denigration of other forms of life or action. Promotion of certain ideas and ways of life merely expands alternatives- people would have the choice to choose their old ways as well.

However, there might be particular abuses involved in particular implementations, but these abuses are not essential to interventions (A)-(C). For example, in WWII, Native Americans were often forced into being scouts and snipers by their commanders due to their image as hunters. Alternative options were closed off for them and a new one was promoted in this instance, which would be an abuse of intervention condition A. Still, such abuses are not necessary to implement these interventions. In this scenario, the government advertising the role of scouts and snipers yet keeping other roles available would not be an abuse of power.

4.5 Objections and Replies

Having made a case for thinking that these interventions to which the government would be committed are not necessarily problematic, and in fact might be freedom-enhancing, I will now consider a couple of objections to my argument.

First, one might object as follows: if the exercise of some capability is to be such that it could appear among one's deliberative options, and if the person in question is already strongly committed to a way of life that does not include the exercise of that capability, then wouldn't the government have to find some way of unfixing this person's commitment to their way of life, which precludes the exercise of the capability? And wouldn't this necessarily involve denigration of that way of life, whether explicit or implicit? And, if this is so, then it seems like the capabilities approach is at odds with political liberalism because it would require the presentation of a robust picture of the good living that precludes others.

In response to this objection, I would argue that, although liberty and freedom of choice are vital to a society, some measure of restrictions must be made if any form of paternalism is to take place. Any bill of rights or constitutional document is in a way paternalistic and restrictive in the sense that it establishes that people cannot behave in every way that they wish to behave. So while individual choices are important, one must note that capabilities often require some measure of functioning in order to fully back them up. While the government should not be focused on unfixing one's commitment to their way of life, and allow for as much freedom as possible, it should aim to foster a climate in which people aim to pursue all sorts of capabilities. In other words, though the government may aim to present all choices as attractive, it wouldn't force or denigrate a particular one explicitly. However, that isn't necessary to present more options and choices for people to make. Being strongly committed to one idea does not mean you must be kept unaware of others. Awareness and education are key- informing people of all options does not contain denigration, but instead enhances people's freedoms. In the case of Aristotle, public programs on the importance of voting and civic engagement would not restrict his freedom, but merely expand on his opportunities to choose.

A second concern arises when we consider how the government might try to meet condition (1) in circumstances in which there really aren't similar enough people exercising the

capability in question. How could the government ensure (1) in such circumstances? It seems like this would have to involve deception of some kind, whether through coercing actual people to engage in the relevant activities, or through AI- and other kinds of generated propaganda. But this, too, would be deeply problematic. Governments should not coerce people, or deceive them. In response, I would argue that, if there are circumstances in which there aren't people exercising the capability in question, then they should find ways to make aware of that option to the people in question. Even so, perception is key- certain groups may be disinclined to take certain capabilities, but having the awareness of that option through the conditions described is enough. For example, take the scenario of racial minorities being underrepresented in certain job opportunities. Promotion of opportunities for them does not mean that deception is required. Once more people are available to ensure condition (1), then that capability can be further fulfilled.

5. Conclusion

We saw in this paper how theorists have envisioned in the capabilities approach a new way of thinking about political justice – what governments should aim to do for their citizens. In this literature, however, there is a real question about whether the capabilities view can be made compatible with other political commitments, like a commitment to political liberalism. The capabilities framework that so many theorists deliberate about is therefore riddled with controversy, such as Terlazzo's disagreement with Nussbaum's view on the compatibility of political liberalism.

In this paper, I express multiple perspectives of the capabilities view and critiques of those viewpoints, and expand on issues I take with those particular viewpoints. I then set my own terms for capabilities by suggesting a two-part condition in which perception of certain attributes determines possession of deliberative options which in turn determines possession of a capability. I also delve into what it would entail in order to fulfill them through paternalism. Taking into consideration different theorists' accounts, I set a framework that describes the capabilities view in necessary conditions and deliberative options. Through such reasoning, I address the criticisms labeled upon the seeming contrast between political liberalism and the capabilities view, coming to a newly revised structure that strikes a balance between preserving freedom and acknowledging the critical role of external support in achieving political justice.

Works Cited

- Claassen, Rutger. "Making Capability Lists: Philosophy Versus Democracy." *Journal of Human Development and Capabilities* 16, no. 1 (2015): 1-15. Accessed August 14, 2024.
- Terlazzo, Rosa. "How Politically Liberal Should the Capability Approach Want to Be?" *Journal of Political Philosophy* 25, no. 3 (2017): 287-307. Accessed August 14, 2024.
- Sen, Amartya. "Capabilities Lists and Public Reason." *Journal of Human Development and Capabilities* 15, no. 4 (2014): 473-478. Accessed August 14, 2024.
- Robeyns, Ingrid. "Sen's Capability Approach and Gender Inequality." *Feminist Economics* 9, no. 2 (2003): 61-92. Accessed August 14, 2024.
- Nussbaum, Martha. "Working with and for Animals." In *Animal Rights: Current Debates and New Directions*, edited by Cass R. Sunstein and Martha C. Nussbaum, 1-24. New York: Oxford University Press, 2004. Accessed August 14, 2024.
- Cripps, Elizabeth. "Saving the Polar Bear, Saving the World." *Environmental Ethics* 30, no. 2 (2008): 151-168. Accessed August 14, 2024.
- Robeyns, Ingrid. "Wellbeing, Freedom and Social Justice: The Capability Approach Re-examined." Cambridge: Open Book Publishers, 2018. Accessed August 14, 2024.
- Nussbaum, Martha Craven. "Women and Human Development: The Capabilities Approach." Cambridge: Cambridge University Press, 2000. Accessed August 14, 2024.
- Vihvelin, Kadri. "Free Will Demystified: A Dispositional Account." *Philosophical Topics* 32, no. 1/2 (Spring & Fall 2004): 427-450. University of Arkansas. Accessed August 14, 2024. <https://www.jstor.org/stable/43154446>.
- Kamtekar, Rachana. "Situationism and Virtue Ethics on the Content of Our Character." *Ethics* 114, no. 3 (April 2004): 458-491. Published by The University of Chicago. Accessed August 14, 2024. <https://www.jstor.org/stable/10.1086/381696>.
- Gandjour, Afschin. "Mutual Dependency Between Capabilities and Functionings in Amartya Sen's Capability Approach." *Social Choice and Welfare* 31, no. 2 (August 2008): 345-350. Accessed August 14, 2024. <https://www.jstor.org/stable/41107927>.
- Piñeros Glasscock, Juan S. "Reason in Action in Aristotle: A Reading of EE V.12/ EN VI.12." *Journal of the History of Philosophy* 57, no. 3 (July 2019): 391-417. Johns Hopkins University. Accessed August 14, 2024.

MBRP as an Effective Method to Treat Substance Use Disorders By Arul Muthuvelu

Abstract

This paper addresses the need to evaluate and update the evidence surrounding MBRP's effectiveness in treating substance use disorders (SUDs). Previous systematic reviews have assessed MBRP's efficacy, but with emerging research, there's a demand for a comprehensive overview of recent findings. Here, we aim to provide an updated synthesis of MBRP research, shedding light on its role in SUD treatment. Mindfulness-Based Relapse Prevention (MBRP) has demonstrated significant effectiveness in treating SUDs by enhancing mindfulness skills that help individuals manage cravings and triggers more adaptively. Studies indicate that MBRP reduces relapse rates, improves coping mechanisms, and supports recovery across diverse populations, including those with alcohol dependence. The structured, group-based nature of MBRP, combined with expert facilitation and ongoing support, reinforces these skills and fosters a sustainable path to recovery. This evidence suggests that MBRP is a versatile and impactful intervention for various forms of substance use.

Introduction

Substance Use Disorder (SUD) is a complex and chronic condition characterized by the compulsive use of drugs despite harmful consequences. Individuals with SUD often prioritize drug use over their personal, societal, and health well-being. This disorder involves a pattern of repeated use that leads to significant impairment or distress, affecting various aspects of life including work, relationships, and physical and mental health. The condition is marked by cycles of relapse and remission, where individuals may struggle to maintain abstinence despite their best efforts and the severe negative impacts of continued use. This is a global issue and a burden to governments and has major effects in the global political scene. For instance, in 2016 alone, there were 290,000 people hospitalized for opioid use.

Effective interventions for SUD must address both the behavioral and neural mechanisms underlying addiction.

Among behavioral modification strategies such as extinction training and replacement therapy, Mindfulness-Based Relapse Prevention (MBRP) has emerged as a promising approach. MBRP integrates mindfulness practices into relapse prevention efforts, offering a holistic approach to addiction treatment. MBRP integrates principles of mindfulness with relapse prevention strategies, aiming to enhance awareness of triggers, cravings, and habitual patterns associated with substance use. It emphasizes non-judgmental observation of thoughts and emotions, acceptance of cravings, and development of coping skills to manage high-risk situations. The program typically comprises several sessions led by trained facilitators, focusing on various mindfulness techniques and their application in daily life.

Methods

Search Strategy

To investigate the effectiveness of Mindfulness-Based Relapse Prevention (MBRP) in treating Substance Use Disorders (SUD), a systematic search approach was devised. This involved employing a range of pertinent keywords and phrases to capture relevant literature on the subject. The aim was to identify studies showcasing the efficacy of MBRP as an intervention for various types of substance use disorders.

Keywords Used

Alcohol Use Disorder - This term was included to encompass studies focusing on alcohol dependence and its treatment using MBRP.

Cocaine Use Disorder - Included to capture research on cocaine addiction and the application of MBRP as an intervention.

Opioid Use Disorder - This term was chosen to incorporate studies examining opioid addiction and the efficacy of MBRP in its treatment.

Marijuana Use Disorder - Included to cover research on marijuana dependence and the potential role of MBRP in addressing this substance use disorder.

Substance Use Disorder - Encompassing term to include studies on various types of substance dependencies, ensuring a comprehensive exploration of MBRP's effectiveness across different substances.

Addictive Personality - This term was included to capture literature focusing on personality traits associated with addiction and the potential effectiveness of MBRP in individuals with such traits.

Effectiveness of Mindfulness-Based Relapse Prevention in Treating Substance Use Disorder - Specifically targeting studies evaluating the effectiveness of MBRP in treating substance use disorders, ensuring a focused exploration of the intervention.

Mindfulness-Based Relapse Prevention Effectiveness - A broad term focusing on the overall effectiveness of MBRP as an intervention for preventing relapse in individuals with substance use disorders.

Search Query

Alcohol use disorder OR
cocaine use disorder OR
opioid use disorder OR
marijuana use disorder OR
alcohol use disorder OR
substance use disorder OR
addictive personality OR
effectiveness of mindfulness based relapse prevention in treating substance use disorder
AND Mindfulness Based Relapse Prevention effectiveness

The search query was formulated using Boolean operators to combine different aspects related to substance use disorders and the effectiveness of MBRP.

The search yielded a total of 35 papers and reports. These documents were screened based on relevance to the study's focus on MBRP as a treatment for substance use disorders. Subsequently, a selection of studies were chosen for analysis based on their direct alignment with the research objectives and their significant contributions to understanding the efficacy of MBRP in addressing substance use disorders. We selected 9 papers. The other reports were either not relevant or we couldn't access them online without having to pay.

Opioid Drugs and How They Work

Opioids are analgesic drugs that bind to the opioid receptors within the brain (Mu, delta, kappa receptors) and include such drugs as: Heroin, Fentanyl, Morphine, Codeine, Methadone, etc. These receptors are part of the G-protein coupled receptor family, which initiates a cascade of intracellular events upon activation. When an opioid binds to these receptors, it inhibits the release of neurotransmitters by reducing the influx of calcium ions into the nerve terminals, thereby decreasing neuronal excitability and transmission of pain signals. This binding also leads to the opening of potassium channels, causing hyperpolarization of the neuron and further inhibition of neuronal firing. The activation of mu receptors, in particular, is responsible for the analgesic and euphoric effects of opioids, as well as their addictive potential. Repeated opioid use can lead to tolerance, where higher doses are required to achieve the same effect, and dependence, characterized by withdrawal symptoms upon cessation. This intricate process underscores the potent impact opioids have on the brain's reward and pain pathways, contributing to their high potential for abuse and addiction.

Evidence from Research Studies

Several studies examined the effectiveness of Mindfulness-Based Relapse Prevention (MBRP) as an intervention for Substance Use Disorders (SUDs), with several focusing on specific adaptations for different substances, such as alcohol.

Across multiple studies, MBRP was shown to be a promising intervention for general SUDs. Participants in the treatment groups received an 8-week course of MBRP, consisting of mindfulness meditation practices, cognitive-behavioral relapse prevention skills, and practical applications of mindfulness in daily life. The sessions were typically delivered in a group format, with each session lasting between 90 minutes to two hours (Killeen, 3). Key components included guided meditations, discussions on personal experiences, and exercises aimed at integrating mindfulness practices into participants' daily lives. Participants were often provided with audio recordings or CDs for home practice to reinforce the skills learned during the sessions. Instructors for these interventions were trained professionals with backgrounds in psychology, social work, and cognitive-behavioral therapy, ensuring a high level of treatment fidelity.

Alcohol Dependence

One specific study tailored MBRP for individuals with alcohol dependence, known as Mindfulness-Based Relapse Prevention Adapted for Alcohol Dependence (MBRP-A). This adapted intervention combined mindfulness meditation practices with cognitive-behavioral relapse prevention skills, tailored to address the unique challenges faced by individuals with alcohol dependence (Zgierska, 9). Like the general MBRP studies, MBRP-A consisted of eight weekly group sessions, each lasting two hours, led by trained therapists. Participants engaged in both formal (e.g., body scan meditation) and informal (e.g., mindfulness of daily activities) mindfulness practices. The intervention aimed to enhance participants' ability to cope with cravings and high-risk situations related to alcohol use through increased mindfulness and awareness.

Therapeutic Communities and Residential Settings

Another study explored the integration of MBRP with standard treatment provided at a residential facility, demonstrating the adaptability and efficacy of MBRP in different treatment settings. Participants received eight additional MBRP sessions alongside their standard treatment (Davis et al., 39). These sessions aimed to help participants integrate mindfulness practices into their daily lives, identify and cope with common triggers, and develop a non-judgmental awareness of their experiences. The sessions were conducted by research staff and focused on the role of thoughts in the relapse process, emphasizing individual experiences in building coping skills.

Comparison of Results

The results across these studies consistently indicated that MBRP is an effective intervention for reducing relapse rates in individuals with SUDs. Participants reported improved mindfulness skills, increased awareness of triggers and cravings, and enhanced ability to cope with high-risk situations. The adapted MBRP-A for alcohol dependence showed similar positive outcomes, with participants demonstrating better coping strategies and reduced alcohol use (Zgierska et al., 11). The studies also highlighted the importance of trained professionals in delivering the intervention and maintaining treatment fidelity. Overall, MBRP proved to be a versatile and impactful approach for addressing various forms of substance use, offering significant benefits in preventing relapse and promoting recovery.

Several studies have investigated the effectiveness of MBRP in treating SUDs, providing valuable insights into its potential benefits. One study from the University of New Mexico conducted among adult women with SUDs in a residential treatment center, compared the outcomes of participants who received MBRP with those who received Relapse Prevention (RP) therapy. The findings suggested that MBRP participants demonstrated improvements in addiction severity and reduction in substance use days compared to RP participants. Additionally, completion rates were higher among racial and ethnic minority participants in the

MBRP group, highlighting its potential effectiveness across diverse populations (Witkiewitz et al., 5).

Another study by the University of Sao Paulo, conducted in a therapeutic community setting for individuals with SUDs, compared the outcomes of participants who received MBRP alongside treatment as usual (TAU) with those who received only TAU (Donate et al., 12). The results indicated that the MBRP group showed promising reductions in substance use and improvements in various domains compared to the TAU-only group. Moreover, the study adhered to rigorous ethical standards and guidelines, ensuring the reliability and validity of its findings.

Potential Mechanisms of Action

The effectiveness of Mindfulness-Based Relapse Prevention (MBRP) in addressing Substance Use Disorders (SUDs) can be largely attributed to its comprehensive approach in enhancing mindfulness skills. This method equips individuals with the ability to observe their thoughts, feelings, and bodily sensations without immediate judgment or reaction. By cultivating a heightened self-awareness, individuals are better able to recognize cravings and triggers as they arise, allowing them to respond in a more deliberate and adaptive manner rather than resorting to automatic, habitual substance use.

One of the core components of MBRP is the practice of mindfulness meditation, which helps individuals develop a present-moment awareness and acceptance of their experiences. This acceptance reduces the emotional reactivity and stress that often accompany cravings, making it easier to manage them without relapsing. Techniques such as breath meditation, body scan, and urge surfing are employed to help participants observe cravings as transient events that can be experienced without acting on them. This shift in perspective enables individuals to break the cycle of automatic responses to triggers and develop healthier coping mechanisms.

Moreover, MBRP encourages the integration of mindfulness practices into daily life, which fosters ongoing self-regulation and resilience. Participants learn to apply mindfulness techniques in real-world situations, enhancing their ability to handle high-risk scenarios and stressful environments that could otherwise lead to relapse. This practical application of mindfulness ensures that the skills learned in sessions are continuously reinforced and utilized in everyday challenges.

The structured nature of MBRP sessions further contributes to its effectiveness. Typically conducted in a group format over eight weekly sessions, MBRP provides a supportive and collaborative environment where individuals can share experiences and insights. This group setting fosters a sense of community and accountability, which can be particularly beneficial in sustaining motivation and commitment to recovery. Additionally, the presence of trained facilitators ensures that the sessions are guided with expertise and adherence to the therapeutic protocol. Facilitators, who often have backgrounds in psychology or social work, provide personalized feedback and support, enhancing the therapeutic impact of the intervention.

Furthermore, the provision of resources such as guided meditation recordings and home practice assignments helps reinforce the skills learned during sessions. These resources allow individuals to practice mindfulness consistently, which is crucial for the long-term maintenance of recovery. By incorporating daily mindfulness practices, participants are better equipped to handle future challenges and reduce the likelihood of relapse.

Overall, MBRP's emphasis on mindfulness, combined with structured sessions, supportive group dynamics, and expert facilitation, creates a robust framework for addressing SUDs. This holistic approach not only helps individuals manage cravings and triggers more effectively but also empowers them to cultivate a sustainable and resilient recovery path.

Discussion

Mindfulness-Based Relapse Prevention (MBRP) has garnered significant attention as a potential treatment for substance use disorder (SUD), particularly in recent years. Drawing from the information provided, let's delve into a comprehensive discussion on whether MBRP is an effective intervention for individuals struggling with SUD.

Considerations and Limitations:

While the evidence supporting the efficacy of MBRP is promising, certain considerations and limitations should be acknowledged. For instance, the generalizability of findings may vary depending on factors such as participant demographics, treatment setting, and duration of follow-up. Moreover, the absence of long-term follow-up data in some studies limits our understanding of the sustained effects of MBRP beyond the immediate intervention period.

Conclusion

In conclusion, MBRP holds promise as an effective intervention for individuals grappling with substance use disorder. Its integration of mindfulness principles with relapse prevention strategies offers a holistic approach to addressing the complex challenges associated with addiction. While further research is warranted to validate its long-term efficacy and explore its mechanisms of action more comprehensively, the existing evidence suggests that MBRP represents a valuable addition to the arsenal of treatment options for SUDs. By fostering greater self-awareness, acceptance, and coping skills, MBRP empowers individuals on their journey toward recovery and sustained well-being.

Works Cited

1. Davis, J. P., Berry, D., Dumas, T. M., Ritter, E., Smith, D. C., Menard, C., & Roberts, B. W. (2018). Substance use outcomes for mindfulness based relapse prevention are partially mediated by reductions in stress: Results from a randomized trial. *Journal of substance abuse treatment*, 91, 37–48. <https://doi.org/10.1016/j.jsat.2018.05.002>
2. Donate, A. P. G., de Macedo, E. C., Bedendo, A., Félix Júnior, I., Gallo, G. G., Opaleye, E. S., & Noto, A. R. (2023). The effect of mindfulness on decision-making, inhibitory control, and impulsivity of substance use disorder in-treatment patients: A randomized clinical trial. *PloS one*, 18(11), e0293502. <https://doi.org/10.1371/journal.pone.0293502>
3. Grant, S., Colaiaco, B., Motala, A., Shanman, R., Booth, M., Sorbero, M., & Hempel, S. (2017). Mindfulness-based Relapse Prevention for Substance Use Disorders: A Systematic Review and Meta-analysis. *Journal of addiction medicine*, 11(5), 386–396. <https://doi.org/10.1097/ADM.0000000000000338>
4. Killeen, T. K., Baker, N. L., Davis, L. L., Bowen, S., & Brady, K. T. (2023). Efficacy of mindfulness-based relapse prevention in a sample of veterans in a substance use disorder aftercare program: A randomized controlled trial. *Journal of substance use and addiction treatment*, 152, 209116. <https://doi.org/10.1016/j.josat.2023.209116>
5. Roos, C. R., Bowen, S., & Witkiewitz, K. (2017). Baseline patterns of substance use disorder severity and depression and anxiety symptoms moderate the efficacy of mindfulness-based relapse prevention. *Journal of consulting and clinical psychology*, 85(11), 1041–1051. <https://doi.org/10.1037/ccp0000249>
6. Weiss, F., Aslan, A., Zhang, J., Gerchen, M. F., Kiefer, F., & Kirsch, P. (2020). Using mind control to modify cue-reactivity in AUD: the impact of mindfulness-based relapse prevention on real-time fMRI neurofeedback to modify cue-reactivity in alcohol use disorder: a randomized controlled trial. *BMC psychiatry*, 20(1), 309. <https://doi.org/10.1186/s12888-020-02717-7>
7. Witkiewitz, K., Bowen, S., Douglas, H., & Hsu, S. H. (2013). Mindfulness-based relapse prevention for substance craving. *Addictive behaviors*, 38(2), 1563–1571. <https://doi.org/10.1016/j.addbeh.2012.04.001>
8. Witkiewitz, K., Greenfield, B. L., & Bowen, S. (2013). Mindfulness-based relapse prevention with racial and ethnic minority women. *Addictive behaviors*, 38(12), 2821–2824. <https://doi.org/10.1016/j.addbeh.2013.08.018>
9. Zgierska, A. E., Burzinski, C. A., Mundt, M. P., McClintock, A. S., Cox, J., Coe, C. L., Miller, M. M., & Fleming, M. F. (2019). Mindfulness-based relapse prevention for alcohol dependence: Findings from a randomized controlled trial. *Journal of substance abuse treatment*, 100, 8–17. <https://doi.org/10.1016/j.jsat.2019.01.013>

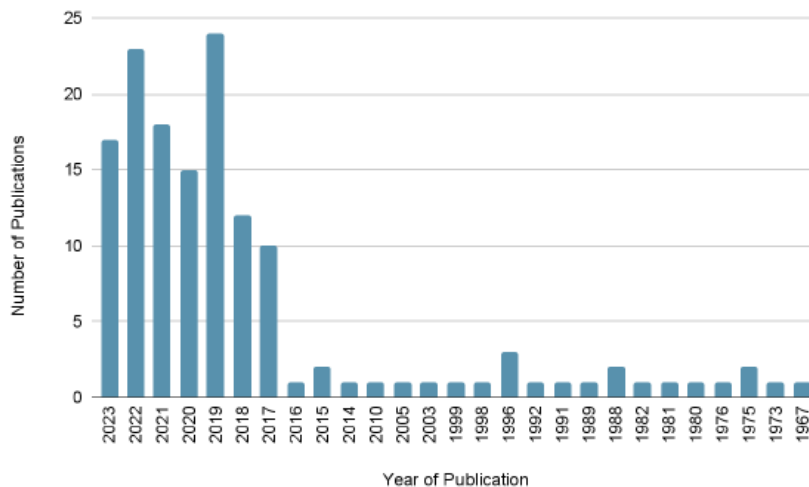
Trends of Down Syndrome Patients Developing Leukemia: A Review of Risk Factors and Early Cancer Detection Using Liquid Biopsy

By Tanisha Arora

Abstract

Down Syndrome (DS), a genetic disorder characterized by the presence of an additional copy of chromosome 21, has been shown to increase the likelihood of developing Leukemia. This study aims to provide evidence on this anomaly as well as share insight on how liquid biopsies can be applied for early detection of cancer. We reviewed multiple studies from the UK and USA and found 194 relevant papers. Based on their relevance, only 31 scientific studies on this topic were shortlisted. In the final systematic review, only four papers were incorporated. The keywords used to search for relevant articles included “liquid biopsies”, “Leukemia”, and “Down Syndrome”. The articles summarized here focus on the biomarkers for Leukemia and present inherited traits that increase the susceptibility of diagnosis for DS patients. Our main finding is that the liquid biopsy approach should be applied to a cohort with an increased risk of developing cancer such as individuals with Down syndrome rather than asymptomatic populations. In addition we highlight the advancements made in using liquid biopsies for the early diagnosis and monitoring patients with blood cancers such as Leukemia.

Chart 1: Relevant PubMed Articles and their Year of Publication



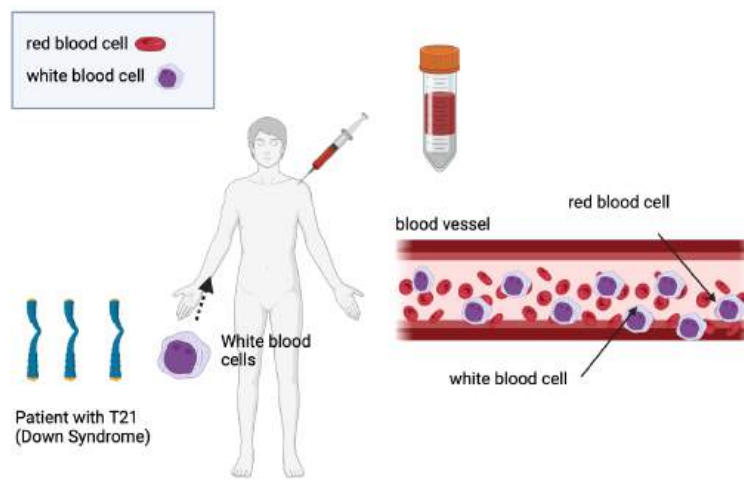
1. Introduction

1.1 Overview

Analytical research concludes children that are diagnosed with DS to be approximately 10 to 30 times more susceptible to developing Leukemia². Discovered in 1958, the presence of Trisomy 21 in DS patients is linked to mitotic errors in cell division, leading to uncontrolled cell

growth and an increased risk of Leukemia. While numerous studies have investigated this connection, further research is necessary to gain a comprehensive understanding. Additionally, in order to reduce the overall number of fatalities, medical professionals should pursue the liquid biopsy method for early detection. This review incorporates data from four articles chosen for their relevance, language, time frame, and country^{5, 6, 7, 8}.

Figure 1: Liquid biopsies to detect Leukemia in individuals with Trisomy 21 (made using Biorender 2024).



1.2 Leukemia and Down Syndrome

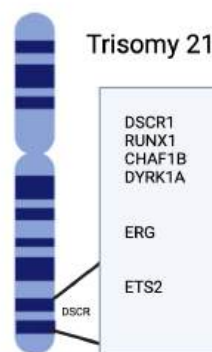
DS is an untreatable genetic disorder caused by an extra pair of chromosome 21. Characteristics associated with this disorder include intellectual decline, facial variations, and developmental disabilities. The runt-related transcription factor 1 (RUNX1) is a gene that regulates multiple cell processes and is present on chromosome 21. As a result of the additional copy of chromosome 21 in DS patients, mutations often occur in this gene which can cause other diseases such as Leukemia³. Leukemia is characterized by the unregulated proliferation of abnormal white blood cells. When there is a genetic mutation in the bone marrow of a patient's blood cells, they tend to increase excessively leading to overcrowding of cells, which obstructs the compulsory cells from performing their jobs. Oftentimes, medical professionals will conduct a complete blood count (CBC) to diagnose patients showing symptoms. During examination of blood cells, doctors will look for either an unusually large amount of white blood cells or few red blood cells.

1.3 Liquid Biopsy

A new approach based on novel methodologies for detecting cancer signals are liquid biopsies. During a liquid biopsy, the clinician will sample a patient's blood to search for cancer biomarkers including circulating tumor cells, (CTCs) which are cells that originate from a tumor or for circulating tumor DNA, (ctDNA) which is DNA derived from the apoptosis of the tumor

cells. Additional biomarkers of this innovative approach include miRNA and exosomes which provide data on the molecular aspects of a tumor for improved management. The steps include using a needle to draw blood followed by separating blood cells from plasma and analyzing them individually to understand the molecular changes taking place. Through such biopsies, medical professionals can provide early detection of Leukemia for individuals at high risk (e.g T21 individuals). By detecting cancer at its earliest stage, they are able to provide the right treatment in order for increased survival rates. Additional benefits of early detection include lower cost and fewer side effects. Since DS patients are already at a high risk for developing cancer, using liquid biopsies to closely screen patients could increase survival rates. However, little research has been done in relevance to diagnosis of hematological cancers through liquid biopsies.

Figure 2: Major genes involved in DS that can cause Leukemia ⁴ (made using Biorender 2024)



1.4 Former Research

A study performed shared the demographics for DS patients that were diagnosed with Leukemia. Data collected from this study revealed the demographic of hispanic male children to have the highest prevalence of Leukemia. Additionally, black children are more susceptible to both DS and Acute Lymphoblastic Leukemia (ALL) compared to white children¹. Finally, this study also summarized that the trends of diagnosis for DS patients compared to healthy patients did not match; between the ages of 2-4 ALL is more common in DS patients, however during the same age in healthy patients Acute Myeloid Leukemia (AML) is more common¹.

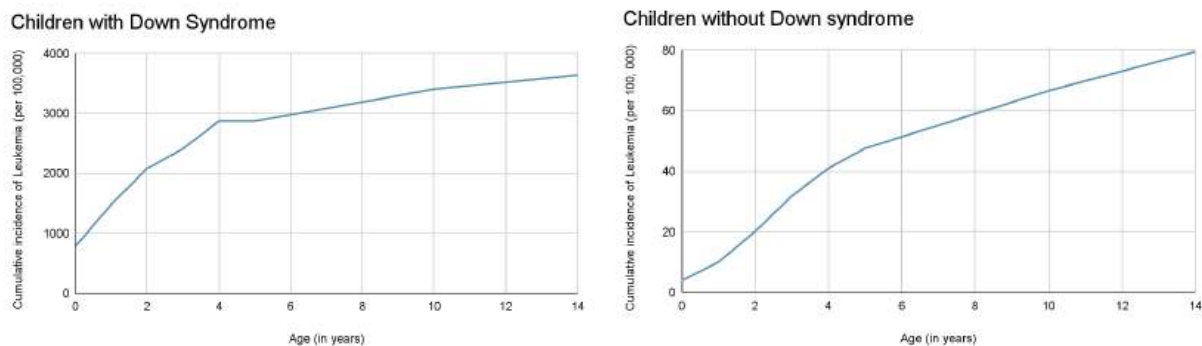
2. Results

According to Winer et al. (2020)⁵, rare gene changes in the reproductive cell which are transferred into the DNA can contribute to ALL in children with DS. Children with DS have an approximately 20-fold increased risk of developing ALL because of Trisomy 21. The presence of this mutation can often prevent B-cell differentiation and increase the frequency of immature stem cells. However, the data collected reported that only ~1% of DS patients developed

Leukemia, which indicates Trisomy 21 is not the only factor responsible for diagnosis. A study was designed to collect data from 73 patients with DS that were diagnosed with Leukemia. In the investigation, researchers noted 143 key genes, as well as three patients who showed significant mutations compared to the rest of the experimental group. Additionally, they discovered three recurring cancer-related genes in these patients including IKZF1, NBN, and RTEL1. For a more comprehensive understanding of the other factors that contribute to development of Leukemia in DS patients aside from chromosome 21 and certain genes, more research is crucial.

According to Marlow et al. (2021)⁶, children with DS have substantially increased risk of multiple health conditions including 10- to 20-fold higher Leukemia incidence than the general population. DS is a strong risk factor for childhood Leukemia, and there is a higher probability of AML as compared to ALL and the following charts illustrate the increased risks. The risk of diagnosis of AML in patients is a 150-fold risk before the age of 5, and is even higher for variants such as AML-M7 at approximately 400- to 600-fold. Previous studies mostly used data from as early as the 1930s, and current data might reflect changes in medical practices, such as increased use of CT scans which are linked to higher Leukemia risk. To update these risk estimates, a study analyzed a cohort of 3.9 million children from recent years. It found that children with Down syndrome continue to have a much higher risk of Leukemia, particularly AML and AML-M7, and that the incidence rates for these conditions remain substantially elevated compared to children without Down syndrome.

Figure 3: Compares susceptibility for Leukemia between children with and without DS (data obtained from Hasaart et al. 2021)



The authors Hasaart et al. (2021)⁷, propose a model in which the increased risk of Leukemia in DS children derives from higher rates of mutation accumulation during fetal development and is further enhanced by changes within the fetal liver niche. This model could possibly be used to understand the steps of leukaemogenesis early in life. The liquid biopsy approach offers the possibility of close monitoring for individuals to allow early diagnosis of leukemia and have a more comprehensive understanding of cancer evolution.

DS patients have a higher susceptibility to diagnosis of certain variants of Leukemia usually in their youth. Trisomy 21 is one of the mutations which when accumulated with other mutations can increase the likelihood of Leukemia. Further mutations are also responsible such

as environmental factors and genetic changes. The study hopes to use data from Leukemia in DS patients to improve the general understanding of Leukemia in healthy patients too. Additional factors responsible for Leukemia include how blood cells are treated during their time in the fetus compared to after birth.

The authors Abdulmawjood et al. (2019)⁸ discussed the role of assessing biomarkers such as circulating tumor cells, circulating tumor DNA, microRNAs, and exosomes in order to understand tumor progression and metastasis. In both children and young adults, hematologic malignancies such as Leukemia are the most common types of cancers. These malignancies all impact blood-associated tissues such as the bone marrow, lymphatic system, and blood cells. Clinical diagnostics of these malignancies rely heavily on the use of bone marrow samples, which is unpleasant for Leukemia patients. Therefore, the usage of minimally invasive methods such as liquid biopsies have the potential to improve the efficacy of diagnostic strategies in Leukemia patients, providing a useful tool for the real-time molecular profiling of patients.

Currently, medical professionals are searching for a noninvasive method of early detection and monitoring of blood cancers to improve survival rates. Traditionally, tissue biopsies have been used to gather this data. However, their invasiveness tends to have many negative effects on the overall well-being of a patient, and in certain circumstances replication is impossible. By focusing on more accessible cancer-derived materials such as CTCs, ctDNA, mRNA, and EXOs, professionals have been starting to perform liquid biopsies in other variants of cancer. This new method known as a liquid biopsy samples a patient's blood to search for potential biomarkers and has substantial benefits in time, cost, and efficiency. Unfortunately, few studies have been conducted on hematological cancers with a focus on Leukemia. Additionally, these biomarkers tend to be found in small concentrations which harms the density and analytical output. More research is necessary to learn more about the risk factors and how liquid biopsy could be used in the clinic to benefit individuals with DS. Some of the clinical applications of liquid biopsy for DS include early detection, identifying new resistance mutations during treatment, and minimal residual disease.

3. Discussion

3.1 Key Findings

Cellular changes during reproduction are the primary cause of DS. This untreatable disorder is often characterized by the presence of Trisomy 21, an extra copy of chromosome 21. Patients diagnosed with DS usually experience both facial and behavioral abnormalities. A study done by the University of California, Davis gathered data from over 3.9 million children born between 1996 and 2016 in multiple healthcare systems ranging from the United States to Ontario, Canada. This statistical analysis reported that 2.8% of pediatric patients with DS are diagnosed with Leukemia, compared to the 0.05% of healthy pediatric patients diagnosed with the same disease¹. From the literature, it was gathered that children with DS have a higher susceptibility to numerous diseases including Leukemia. DS patients are approximated to have a

20-fold increased chance of diagnosis for Leukemia in comparison to healthy children. The higher susceptibility is usually affiliated with ALL and AML. Research has shown that numerous factors are responsible for Leukemia and the increased risk is not just limited to Trisomy 21. Additional responsible factors include cancer-causing genes. Some key genes include IKZF1, NBN, and RTEL1. The studies shared that during the growth of the fetus, certain mutations may occur to allow an increase of cells that promote the development of Leukemia. Currently, research is being performed to improve the knowledge of the factors responsible for Leukemia in DS patients to improve treatments and increase survival rates, however the application of liquid biopsies for early screening of cancer in individuals with T21 is not yet in the clinic.

The standard methods of detection for Leukemia are not optimal because they can result in late stage diagnosis which can lead to low survival rates. The aim should be to offer earlier screening in this High risk Population using liquid biopsy. By adapting the liquid biopsy method to become applicable in blood cancers, early detection rates will significantly increase. This method offers numerous benefits compared to the traditional approaches such as the noninvasiveness, lower cost, and faster results. Regardless of such benefits, medical professionals must figure out a way to modify the procedure to become more targeted to the specialized cells in Leukemia.

3.2 Future Research

Going forward, to reduce the overall number of fatalities, medical professionals should utilize the liquid biopsy method for early detection of Leukemia in DS patients as a high risk population. As a result of the presence of Trisomy 21, patients with DS are already at a heightened risk for diseases such as Leukemia. At the moment, there is no optimal approach for the detection of Leukemia. The standard methods are costly, invasive, and provide limited information. However, incorporating the liquid biopsy method into daily practices for diagnosis can allow for more efficiency, accessibility, reduce the cost, and have less invasiveness. Additional benefits include more specialized patient care to improve treatment plans. To start applying this method, medical professionals must improve the technology used for detection to be able to identify more specialized cells, as well as adapt the traditional protocols.

3.3 Limitations

During the composition of this literature review, several constraints were dealt with. These limitations included the time frame and lack of research on liquid biopsies as a method of detection for blood cancers. To prevent sharing any outdated information, a time constraint was established. Only articles published within the past five years were incorporated. Therefore, some articles sharing key insights were disregarded to maintain the accuracy of this review. An additional restriction dealt with was the shortage of literature published discussing the usage of liquid biopsies to diagnose Leukemia. The insufficient amount of research on the application of liquid biopsies made it difficult to provide a detailed analysis. Regardless, this literature review aimed to share data and offer aspects of medicine that require further research

4. Methods

The articles included in this review are all published within the past five years (2019-2024) to ensure they are state of the art and with most up to date information. Additionally, all of these research papers were written in English. All research included has been done within both the United States of America (USA) and the United Kingdom (UK). Another aspect in selecting papers was ensuring most of the studies had clinical trials and were not systematic reviews. To extract relevant information a variety of articles were searched for novel findings as well as numerical data supporting the thesis. Finally, keywords such as “liquid biopsies” and “Down Syndrome” were also used to make sure the papers included studies that were relevant to the topic of interest. Originally 31 papers were short-listed, however, after reviewing them in regards to the criteria, only four were chosen. The papers that were excluded were due to lack of relevance, time frame, or country of research.

5. Acknowledgements

Dr. Maria Neofytou of Cancer Research UK Cambridge Institute and Cambridge Centre of International Research provided the guidance to compose this literature review.

6. References

Table 1: Detection of Hematological Malignancies

First author, year of publication	Number of patients	Investigating molecular markers	Type of Leukemia	Method of detection
Peleg Winer, 2020	73 children	IKZF1, NBN, RTEL1	ALL	Germline WES
Emily C. Marlow, 2021	3,905,399 children	N/A	ALL, AML	ICD9/10 diagnosis code
K. A. L. Hasaart, 2021	N/A	GATA1, CTCF, EZH2, KANSL1, CRLF2, IL7R	ML-DS, ALL	N/A
Bilal Abdulmawjood, 2019	N/A	NPM1, FLT3	AML, CML, MDS,	LB (CTCs, ctDNA, miRNA, EXOs)

Works Cited

1. <https://health.ucdavis.edu/news/headlines/study-finds-risk-of-leukemia-higher-than-expected-in-children-with-down-syndrome/2021/04#:~:text=It%20found%20that%202.8%25%20of,to%200.05%25%20of%20other%20children>
2. <https://www.sciencedirect.com/science/article/pii/S2468124521000012>
3. <https://www.news-medical.net/news/20230310/Scientists-identify-the-cause-of-leukemia-in-people-with-trisomy-21.aspx#:~:text=Scientists%20led%20by%20the%20Department,that%20seems%20to%20be%20 responsible>
4. <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2021.636633/full>
5. P. Winer, I.S. Muskens, K.M. Walsh, A. Vora, A.V. Moorman, J.L. Weimels, I. Roberts, A. Roy, A.J. de Smith, Germline variants in predisposition genes in children with Down syndrome and acute lymphoblastic leukemia, *Blood Advances*, **4(4)**, 672-675 (2020).
6. E.C. Marlow, J. Ducore, M.L. Kwan, S.Y. Cheng, E.J.A. Bowles, R.T. Greenlee, J.D. Pole, A.K. Rahm, N.K Stout, S. Weinmann, R. Smith-Bindman, D.L Miglioretti, Leukemia Risk in a Cohort of 3.9 Million Children With and Without Down Syndrome, U.S. Department of Health and Human Services, **e3**, 172-180 (2021).
7. K.A.L Hasaart, E.J.M. Bertrums, F. Manders, B.F. Goemans, R.v. Boxtel, Increased risk of leukemia in children with Down syndrome: a somatic evolutionary review, Cambridge University Press, **e5**, 23 (2021).
8. B. Abdulmawjood, C. Roma-Rodrigues, A. R. Fernandes, P. V. Baptista, Liquid biopsies in myeloid malignancies, *Cancer Drug Resistance*, **2(4)**, 1044-1061 (2019).

The Utility of Art in Depicting the Contrast Between the Women of Sparta and Athens

By Vivienne Tang

Abstract

Were women portrayed in art during the 5th and 6th centuries? If so, how did the location of the artwork alter the depiction of them? This paper argues that analyzing historical artworks featuring women from Athens and Sparta allows for a comparison of women's socioeconomic status in these two city states. Women in Sparta and Athens lived very different lifestyles because of their city's opposing belief systems. Spartan women were known to have many more rights than the other women of their time, such as Athenians. Art plays a significant role in history, even when compared to texts, because of its visual depictions with meaning. Art can be interpreted in a variety of ways and reflects the culture of where it was created. In this paper, I will analyze artwork to learn about the role of women in Athens and Sparta. The depiction of Spartan women in art proved to be a reflection of the higher status women held in that city compared to Athens.

Introduction

Athens and Sparta were two city-states located 132 miles apart in what is now Greece, and both were especially prominent from the sixth to fourth century BCE. Athens was part of mainland Greece, while Sparta was in the southern Peloponnese. Belying their relative physical proximity, these two cultures held opposed belief systems. Athenians were known for placing a high value on education and culture (Vanderpool and Ehrlich). These cultural beliefs led them to treat men and women differently in terms of rights, education, and expectations. Boys of affluent families were always educated in academic subjects, as well as in extracurriculars. Conversely, girls were not commonly educated in any way that was not part of life at home, such as cooking, raising children, and weaving. Some Greek women did receive an education, but this would only happen if they were incredibly wealthy (Cole 129-155). Athens also prided itself on being a democracy, which allowed equality for a small group of men ("What Was Most Valuable to Athenians?"). Only free, adult men, which made up ten percent of all of Athens' population, were allowed to vote ("The Government of Ancient Athens"). On the other hand, society considered higher class Spartan women as men's equals and did not perform the typical arts of Athenian women like weaving or spinning. Instead, that would be left to helots, while the women spent their time taking care of their property and participating in sports. Sparta had a highly esteemed military which was the most important part of their livelihood ("Sparta"). At the age of seven, both boys and girls would learn about humanities, mathematics, survival skills, and athletics ("The Two Faces of Greece"). Sparta was an oligarchy with two kings, but was also an aristocracy. Throughout ancient Greece, people primarily worshiped the same set of gods and goddesses, such as Athena, Zeus, and Ares. However, each city state had specific deities they worshiped, depending on the values of that place. For instance, Athens has a founding myth surrounding Athena and Poseidon. In this story, the two fight over the patronage of the unnamed

land, soon to be Athens. They decided that each of them would give a gift to Cecrops, the king of Athens, and he would decide whose gift he liked better. Poseidon gave him a horse and Athena offered the first domesticated olive tree. Cecrops chose the olive tree, a symbol of peace like Athena (Miller). This central belief of peace reflects the significance of it in Athens.

A major event that involved both of the two city-states was the Peloponnesian War. The Peloponnesian War was a war fought between Athens and Sparta from 431 to 404 BC. This war altered ancient Greece because Athens, the original power of Greece, was losing strength to Sparta, a growing city state. Historians periodize the Peloponnesian War into three eras: the Ten Years War, the Peace of Nicias, and the Ionian War ("Peloponnesian War"). During the Ten Years War, Spartan king Archidamus was able to invade Athens multiple times with the Peloponnesian army. At the same time, Athens used the navy of the Delian League to attack the Peloponnesian coast. From this, the Peace of Nicias, a peace treaty between Sparta and Athens, was signed. There was peace between them from 421 BC to 413 BC. Soon after, the Sicilian Expedition caused Athens to lose most of their navy while trying to capture Syracuse, an ally of Sparta. This event once again revived the war between Sparta and Athens, marking the beginning of the last phase, the Ionian War. This time, the Persian Empire was supporting Sparta with a goal of gaining back Greek cities of Asia minor. In the end, Athens was overtaken by Sparta.

One should keep in mind that most of the sources documenting the Peloponnesian War were written from the Athenian's perspective because Sparta did not produce lasting records. The main source describing the war is *The History of the Peloponnesian War* by Thucydides. Thucydides was an Athenian historian but also a general in the Athenian army, so he recounted the experiences he had during the war. His work was split up into eight books, based on the different parts of the war (Bonner 73-82). In book two, Thucydides characterized Brasidas, a highly accomplished Spartan officer, as "...brave, vigilant, and active... In a word, the distinguishing excellencies of both the Spartan and Athenian characters seem to have been united in this Brasidas (Edwards and Bushnell 62)." While Thucydides was an Athenian, he clearly showed a liking to Brasidas, a Spartan. He described Spartans as physically strong and resistant, but also as improper and conceited. Alternatively, Athenians were depicted as gentle and intelligent. In his words, Brasidas represented the positive characteristics of both Athenians and Spartans. Thucydides only praised the Athenians, whereas he both praised and criticized Spartans. His bias for Athens is shown throughout his work, so one can not uncritically accept what he writes as the truth.

This paper illustrates the differences between the art depicting the women of Sparta and Athens. It uses art because it has been a common outlet for peoples' beliefs for years. Unlike texts, which are typically straight to the point, artwork has the ability to be understood in multiple ways, making it more intriguing. It can carry hidden values or meanings that are not obvious to the eye. Today, most Spartan writings no longer exist, so most of the texts describing it were made from the Athenians perspective. Art can act as a supplement to the history that is written down, most of which was written by Thucydides. The pieces of art selected all fit a common theme of being produced from the late seventh to fifth century BCE in Sparta or

Athens. There are also three specific themes the artworks are split up into: goddesses, women as decorations, and women's life.

For the purpose of this paper, it is important to define the key vocabulary related to the topic of art in Sparta and Athens. In this paper, art is defined as an application of creativity to depict something visual. A figurine is a person or animal who is formed into a small sculpture. A stela is a carved stone slab or pillar used to commemorate someone. A relief is a type of sculpture that is one-dimensional. A lebes gamikos or nuptial lebes is an ancient Greek marriage vase. Red-figure is an art technique wherein the background of a vessel would be covered in black slip, leaving the figures bold in red (McKeown and Sneed). Helots is the name for men and women who were forced into work in Sparta almost like slaves. A lekythos is an oil flask that is used in many places including the kitchen, bathroom, and during funerals.

Section 1: Goddesses

Figure 1 is a small Laconian, a lead figurine that was created sometime during the late seventh century to early sixth century BCE. It was found in the Sanctuary of Artemis Orthia in Sparta, along with over one-hundred thousand similar figurines. The body of the lead figure is front-facing, while the head is turned completely to the side. Both the arms and legs are broken off. The dress acts as the focal point of the artwork, as it is the largest part and the most detailed. It falls down to the goddess' ankles and stretches out besides both of her shoulders. The whole dress is covered in lines that begin to flare out by the shoulders and almost look like wings. The goddess' face is also a prominent part of the figure, specifically her nose. Her hair sticks straight out to the left, while her mouth is dropped open. Lastly, she has a crown-shaped headpiece that is cracked in the middle.

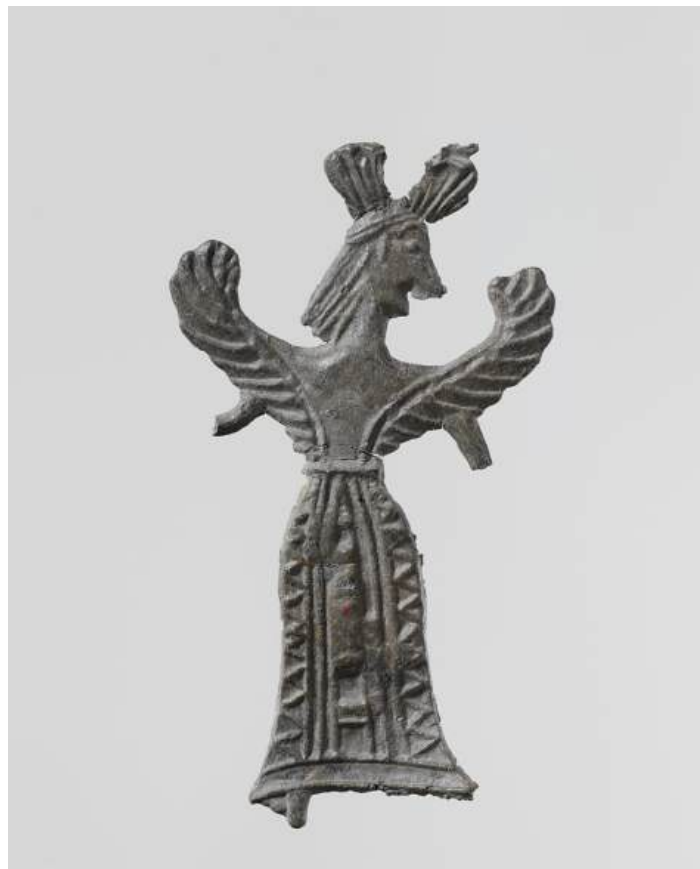


Figure 1: Lead Figure of a Winged Goddess, 6th century, Laconian

This figure of a goddess was most likely supposed to portray Artemis Orthia. It is assumed to be this goddess because the figure shares many features of the “winged Artemis.” In the *Journal of Hellenic Studies*, M.S. Thomson wrote about how the early type of Artemis is represented stranding tall with wings and an animal or bird in her hands. He explained that this

identification of Artemis is often discussed because there are various depictions of her (Thompson 29). Artemis Orthia was the Spartan goddess of wild animals, hunting, vegetation, chastity, and childbirth. Artemis was a figure of female strength by participating in predominantly male activities like battle and sports (“Let's Travel Back in Time: A Look at the Sanctuary through the Ages...”). Her positioning with her arms outwards and firm stance show that she is very powerful. Her arms out to her sides resemble a bird about to fly away, which has commonly been correlated with freedom. This reflection of Artemis Orthia with wings symbolizes her “freedom” through her ability to fly away on her own. As a woman, she was strong and independent, like the Spartan women who had the right to do whatever they pleased. Artemis Orthia’s large role in Sparta reflected the beliefs of citizens when it came to women in physical activities. Artemis being portrayed with symbols of athleticism and freedom show that Sparta encouraged women to engage in activities that Athenian women were not allowed to do. Xenophon was an Athenian military leader and author, who is famous for his philosophical thinking. While he was born as an Athenian, he ended up being exiled from Athens because he campaigned with the Thracians and Spartans (“Xenophon”). In Xenophon’s *Constitution of Lacedaemonians*, he refers to girls’ exercise programs and trials of strength for women, both of which were common practice (Spears 32). A prominent leader writing this confirms the accuracy of women actually being allowed to play sports.



Figure 2: Temple-Relief, 438-432 BCE, Attic

Figure 2 is part of the parthenon sculptures in Athens made from 442-438 BC. It was part of the East frieze before being placed in the British Museum. The relief illustrates the procession of the Panathenaic festival, which is the commemoration of the birthday of Athena. The relief is a large marble rectangle with a large chip off of the top right corner. It shows five girls walking in single file wearing loose, draped dresses that almost make them resemble traditional Greek columns. Two of them are holding long items on a wide platform used to burn incense. The others are carrying bowls and jugs used for rituals. The heads had previously broken off and they now sit in the Acropolis Museum and the Louvre ("Parthenon. East Frieze. Block 8 (VIII). Fragments").

Athena was the goddess of wisdom and courage, but also the figure Athens was named after. In C.J. Herington's words, from *Greece & Rome*, "Athena protects her city... As a physical witness to this belief there began to rise...an Athena-temple greater than any hitherto seen on the Acropolis (Herrington 61-73)." In this quote, it is clearly stated that Athena was the protector of Athens, which made her the most important figure of the city. Athena had great power in Athens by representing the values of Athens. The temple that Herington discussed is the Parthenon. The women in this relief are all following in line to give gifts to Athena, displaying their devotion to her. They are all looking completely forward with great concentration, without any hesitation. This relief shows a clear depiction of women both physically and mentally looking up to Athena. Greek mythology, unlike many religions at the time, portrayed gods and goddesses with human bodies. This meant that the Greeks thought that human bodies could reach perfection by looking like gods. There are often sculptures of them with idealized features including a muscular body. Not only were sculptures of mythology commonly idealized, but also just any sculptures in general. In this frieze, the women are all around the size as Athena Parthenos, a sculpture also on the Parthenon ("Reconstruction Drawing of Cult Statue of Athena Parthenos in Cella of Parthenon"). Seeing as the women in this frieze share similar physical traits with Athena, the artist was presumably trying to show how women can share the same characteristics with Athena. The girls are also all almost identical in size, clothing, and position. In art, size and proportion is often purposeful to show that the character or item is significant. Here, the five girls are equal in size, so the artist may have been trying to show that they are equal in importance. The creator may have been specifically showing how each woman who is faithful, no matter their status, is equal. The women in this are worshipping Athena as someone who they look up to, which reflects how the female standards were highly based off of her.

Since Athena is the patron goddess of Athens, she was commonly portrayed throughout art, especially from the Acropolis, which was dedicated to her. Most of the art on the Acropolis was made of either bronze or marble. The major statue of Athena in Athens was the Athena Parthenos, created by the sculpture Phidias, which was unfortunately destroyed when the Persians captured Athens (Stewart 1-2). Later, Athenians ended up creating another that is now lost. Athena is often referred to as Polias, meaning "protector" of Athens, a very influential and powerful role (Tomlinson 1-2). In contrast, Athena is not typically seen in Spartan art because they did not worship her as much as other goddesses. Sparta's primary goddess was Artemis

Orthia, who is often shown in Spartan art. Although they did not often worship Athena, the Spartans created the Sanctuary of Athena Chalkioikos on the Acropolis of Sparta. This is the oldest sanctuary in Sparta and was once a gathering place for military men ("Sanctuary of Athena Chalkioikos"). This temple was also known as the "Bronze House" because the walls were decorated in bronze plaques, and her cult statue was made of bronze (Cosette). However, this temple no longer exists, so historians are not completely sure what it looked like (Tod 609-614). The fact that the Spartans even had a sanctuary made for Athena reflects how they did in fact worship a considerable amount. One similarity that historians do know that the portrayal of Athena in Sparta and Athens shared was the material. While both Athens and Sparta were primarily focused on Athena and Artemis, there were also a variety of other mythological figures shown in artwork.



Figure 3: Vessel, 410 BCE, Attic



Figure 5: Vessel, 410 BCE, Attic

Figure 3, 4, and 5 are part of a hydria, an ancient Greek water jar, made in 410 BC Athens. The vessel was made by the Meidias Painter, who painted scenes of women in the red-figure technique (“Medias Painter”). In Figure 3, there are only four characters shown with three of them being large and the other one miniature. Himeros is the small, winged person holding out his arms to the right. He is illustrated nude, except for a thin headpiece that goes around his forehead. On his right is Phaon playing a lyre while sitting on a rock and looking slightly to his right. He has a similar shaped headpiece on and a cloth laying on one of his legs. In the middle, Demonassa is holding a thin stick that has been identified as a gold crown. It is colored differently than the rest of the vessel, excluding her bracelet, necklace, and earrings. She has three bracelets on each arm, which look like bangles. Her earrings are single, gold spheres. Her necklace resembles a choker with seven golden spheres. Demonassa is also wearing a similar headpiece as Phaon and Himeros. The upper part of her dress is striped and thin, revealing her chest. The lower part is much thicker and is covered in stars.

In the story of Phaon, he is an old ferryman who takes Aphrodite, dressed up as an old woman, across the Straits of Lesbos for free. In return, Aphrodite turns Phaon into a young man for eternity (Farnell 139-146). In this vessel, he is shown with gold jewelry and is surrounded by mythological figures, which shows how different his life is after becoming a young man. Before, Phaon was just a ferryman who would never be able to spend time with gods and goddesses. When his looks changed, he also became wealthy enough to afford a gold crown and had everything that he could possibly want. This illustrates the belief that beauty correlating with success and happiness was a common belief throughout Athens. This also explains why Aphrodite acted as such a large character in Athens and actually had a connection to every part of a person’s life including marriage, fertility, politics, and even military power (“Worshiping Aphrodite”). While Aphrodite did have a connection to military power in Athens, people were not focused on this aspect of her, unlike those in Sparta. The mythology that the Athenians believed in usually were not focused on physical power, and instead on other topics such as wisdom. Aphrodite was mainly worshiped for her beauty and femininity. This vessel demonstrates how beauty was considered extremely important in Athens.



Figure 6: Stela, 330-317 BCE, Attic

Section 2: Women as Decorations

Figure 6 is a marble stela of a young girl holding a handheld mirror. The stela is in close to perfect condition, considering it was made in 330 to 317 BC Athens and only part of the mirror is broken off. The girl is standing in a rectangular frame with a triangle part at the top. Her hair is put into ten thick braids, all stretching from her forehead to the back of her head. She has a neutral facial expression that does not show any specific emotion. The mirror she is holding is relatively small and round with no details. Her dress is draped across her and actually shows part of her left knee. Her knee is slightly visible because the dress is meant to look like a thin material, but the rest of her legs are not shown.

The curator of this piece, who is from the British Museum, believes that the individual highly resembles the famous statue of Korinna. Korinna was a lyric poetess who usually wrote about local mythology in the Boeotian dialect (“Corinna”). Unlike many other poets in Athens, Korinna was a woman, and, furthermore, she was very well-known for her excellence in writing. While most girls in Athens were only educated in household chores, extremely wealthy ones could learn to read and write (Howitt-Marshall). This stela was most likely used for a young girl’s funeral that was not Korinna, since the curator says that it only “resembles” her. Like Korinna, the daughter may have enjoyed writing, so the family may have wanted the stela of their daughter to look like her to seal her accomplishments in others’ minds. The Fitzwilliam Museum wrote, “The oracle of Apollo at Delphi demanded of the ancient Greeks, ‘Know thyself’, and mirrors have often been used as symbols of wisdom and self-knowledge (“Mirrors”).” The girl may have been holding a mirror to present the daughter’s intelligence and what kind of person she was. Wisdom is typically correlated with education, so although Athenian girls were typically not educated, this stela shows how wisdom held a large value in Athens and how women could still be praised for it (Kapparis).

Figure 7 is the support stick of a bronze Athenian mirror that was made from 490-460 BC. There is a woman in the middle, who is perhaps Aphrodite, standing on a pedestal. She is holding her loose dress with her left hand, and, in her right hand, she is holding a dove. She is standing completely straight, while looking in front of her with a blank expression. On top of her head, there is a triangle and half-oval-shaped items. The two ends of the triangle hold two erotes, winged gods that often accompanied Aphrodite, holding one of their arms on Aphrodite’s shoulder.



Figure 7: Mirror, 490-460 BCE, Attic

Aphrodite was typically associated with the appearance of women and how compliance with the beauty standards was a vital part of womanhood. Agnew wrote, “For the ancient Greeks, a beautiful body was considered to be direct evidence of having a beautiful mind. If you were considered beautiful then you were automatically a good person, and the Greeks actually had a word for this. ‘Kaloskagathos’ meant gorgeous to look at, insinuating a good person... Ancient Greek females also had body ideals to follow... Women looked up to Aphrodite... and depicted her with a round face, large breasts and a pear-shaped body. This then became the beauty ideal for Greek women (Agnew).” Aphrodite played a large role in Greeks’ idea of beauty, which caused many women to look up to her as an idol. Abiding by the Athenian beauty standards determined whether others believed that someone was a good person or not. This illustration shows Aphrodite holding out a dove in her right hand. In general, doves have been seen as a

symbol of femininity and fertility, which is also part of what Aphrodite represented (Willette). Aphrodite having her hand out with a dove is a representation of her offering fertility to the woman who is using the mirror. This shows how becoming a mother was one of the most important aspects of Athenian life for women.

Figure 8 is a Laconian bronze mirror with a nude girl as the support stick. It was created sometime during the second half of the sixth century BC. The actual mirror part is circular with a rim that has clouds around it. The bronze material is very rusted, causing the mirror and handle to be a greenish-brown color. The girl below the mirror is nude, except for some jewelry including a necklace, headpiece, and string across the body. The headpiece has two strips on top of each other that are made up of ovals and two flower-shaped parts on the sides. The necklace has a thick string with one large gem in the middle. The girl is standing on a curled up lion and on both of her sides, there are griffins also holding up



Figure 8: Bronze Mirror with a Support in the Form of a Nude Girl, 6th Century BCE, Laconian

the mirror. In her left hand is a pomegranate.

The fact that the girl is nude and has many animals surrounding her identifies her as probably being the Mistress of Animals, an ancient Near-Eastern deity. The Mistress of animals was primarily known as a fertility goddess, but also similar in some ways to both Aphrodite and

Artemis. Like Aphrodite, she was a symbol of childbirth. She shared her love for slaying animals and using bows with Artemis (“Artemis and Aphrodite”). Day, from the University of Winnipeg, writes:

“The clearer aspect of Anat’s role as mistress of animals is predatory: she is a hunter. The most sustained and nuanced depiction of her association with the hunt is in the Aqhat story...Anat covets the hunting bow and offers Aqhat riches and immortality in exchange for it. He refuses to give Anat the hunting bow, and she vows revenge (Day 181).”

Anat was a warrior goddess and prominent character in the Canaanite mythological texts, a religious writing from an area called Canaan. The Canaanites were an ethnic group from the Bronze Age in today’s Palestine (LaRocca-Pitts). This story is important because it makes the Mistress of Animals’ personality clear. Instead of abiding by the typical female expectations of staying at home to take care of children, she decided to become a hunter, a more “masculine” job. Knowing Sparta highly valued having a commanding military, the Mistress of Animals reflected Spartan principles of how a woman should act. In this piece of art, she has griffins and lions on her shoulders, both animals who are known to be predators and skilled hunters. Their being on her shoulders is a symbol of their similarity, and her position in the middle makes her seem like she is commanding them. As part of a mirror, she is supposed to be someone that reflects the ideal woman, which shows how women were encouraged to have characteristics that other states would have considered improper.

These two mirrors are very similar in design, with the only major difference being the style of the two deities. The Athenian one is much more realistic, while the Laconian one is highly abstract. This is portrayed through the detail shown on the Athenian goddess. Both her



Figure 9: Lebes Gamikos, 4th Century BCE, Athenian

hair and dress are stylized in a way that makes the sculpture life-like. Contrastingly, the Spartan goddess does not have as much detail besides her hair and is actually nude. The positioning of both of the women are almost the same. They have their right hands outward holding something and their other arm at their side. The rest of their bodies are stiff with two legs parallel to each other. The Athenian goddess has a neutral facial expression and the Laconian one has a slight smile. In terms of pathos, the Spartan figure convinces the viewer of an accepting feeling, while the Athenian one reflects a more conservative perspective. While the women are depicting the same thing, they are shown in two very contrasting styles, which reflects the differences in meaning.

Section 3: Women's Lives

Figure 9 is a lebes gamikos, a special kind of wedding vase, that was made in 350 BC. It was discovered on Thasos Island but is made in the Attic style. The vase itself is in an apple-like shape, covered in the red-figure technique. The top of the vase has two arches and a balloon-shaped cover in the middle. It depicts five girls and one winged figure. The winged individual, Eros, also known as Cupid, is standing on a rectangular object the woman second to the right is holding. All of the women are turned to the right, excluding the woman on the very right. They all also have very neutral facial expressions. The woman on the very right is wearing a dress with a highly-detailed scarf, while the rest of the women are wearing very plain dresses that all look relatively similar. The second woman to the right is carrying the rectangular item and a large vase. The woman in the middle has a large, square-shaped item that is covered in cloth above her head. The second woman to the left has a miniature square item and a larger one in her other hand connected to a scarf. The woman to the very left has a cloth in her hands.

This scene is meant to depict a procession of women giving gifts to a new bride, also known as “epaulia.” Giving gifts was a very common practice for Athenians, which was not only done for weddings, but also for worshiping gods. This shows how large of an event marriage ceremonies were and how they carried great importance. In this red-figure, none of the women actually look happy, even though they are celebrating a joyful thing. The fact that all of the women are waiting in line with different items could represent how marriage is a reality for all of them. The items that each woman is holding could be symbols of what they will have in their marriage. For instance, the woman holding a platform with Eros on it may experience true love, since Eros is the god of love (“Eros/Amors/Cupid”). She is also carrying a vase, which has been seen as a maternity symbol, meaning that her children will play a large role in her life (“Vase”). Another interpretation is that each woman is holding the qualities that create a good marriage. This meant that, instead of the woman holding Eros having true love, the artist could have been stating that love is a fundamental element of marriage. The women behind her are holding tapestries, extremely expensive decorations only for the wealthy, representing how they wished for prosperity (“History of Tapestries from Medieval to Modern Times”). The Athenians’ wedding customs represented how marriage was expected for women and specific standards as to what created a good marriage (Denis).

Figure 10 is a bronze sculpture of a Laconian woman made sometime from 520 to 500 BC. The figure was most likely part of either a vessel or a utensil based off of the feet positioning (“Bronze Figure of a Running Girl”). The woman's head is turned ninety degrees to the left, showcasing her long, braided hair that is out of her face. Both her legs and arms are bent into a running posture, while she is holding her dress with her right hand. Her dress is small with only one sleeve, revealing her chest and lower body. Her whole body seems very muscular, especially her legs.

Her legs are proportionally larger than the rest of the body, which was probably done intentionally by the artist so it would catch the viewer’s eye. It is hinting to the observer that her legs are a meaningful part of the body. In this case, it pushes the observer to realize that the woman is an athlete, specifically some sort of runner. Not only were Laconian women doing sports, but they were doing sports that, in other places, would have been solely for men and gave women muscles, similar to those in the bronze figure above (“Ancient Greek Girl Athlete”). This



Figure 10: Figure, 520-500 BCE, Laconian

figure is also not covering up as much as other women in the sixth century would have. The primary reason why women were so encouraged to partake in exercise is because Spartans believed that having strong babies was incredibly important and that having an athletic mother would lead to this. Lycurgus, a lawgiver of Sparta, actually stated that the birth of children was a woman’s number one priority, so they should exercise no less than any man (Plutarch, Christensen 193-255). This Spartan woman represents how Spartans supported activity and physical strength in everyone.

Figure 11 is an Attic, terracotta lekythos, an oil flask, that was made by the Amasis Painter from 550 to 530 BC. It has a symmetrical hourglass shape covered in black-figure. Lekythoi are typically used to hold oil, however, this one was most likely used for funerary purposes. It has two specific regions: one strip on the top around the mouth and one thicker strip in the middle of

the lekythos. The two areas are separated by four straight lines and two dotted lines in between them. As shown in Figure 12, there are thirteen people surrounding the mouth. Four of them are



Figure 12: Terracotta Lekythos, 550-530 BCE, Attic

holding hands and are possibly women dancing because you can see that their limbs are bent. The figure in front of them and the one directly opposite from them looks like they are serving the person in the middle, since they are holding a staff shaped object. The next person also seems like they are serving the person in the middle, but they are speaking. The character in middle, across from the handle, is sitting and stands out compared to everyone else due to their size. Because of this, one can infer that this person is a vital part of this lekythos' story. Considering that this was a funerary object, the figure is either the person who has passed away or a goddess to protect them. She is holding a circular object which, it seems, she will give to the figure on the right. The top four people on the right side also look like they are dancing, but without walking around or moving their legs.

For the second area, depicted in Figure 13, the first full figure on the left seems to be holding two tools: one oval-shaped object and one long, thin object below. The next figure looks like they are holding wool and stretching it out into string. The two people after them are standing in front of a square-shaped object with many vertical lines inside of it. These two

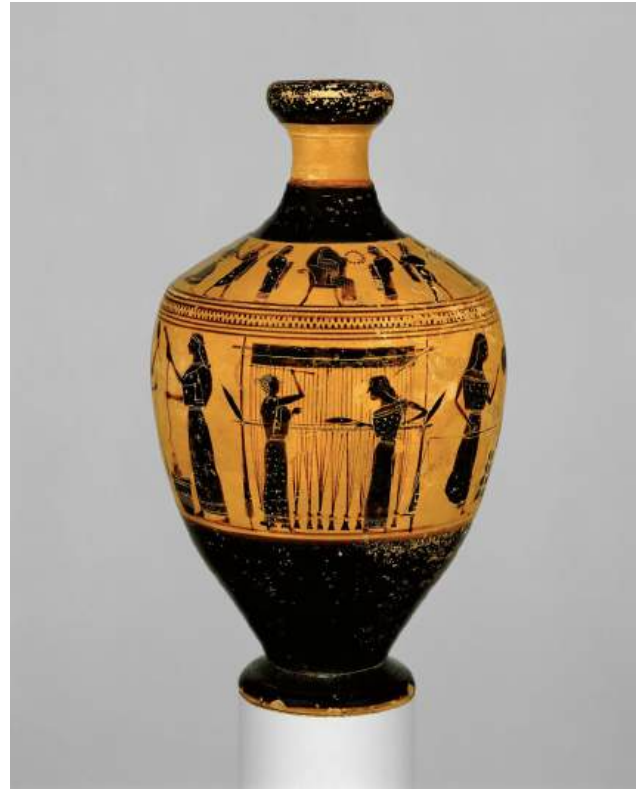


Figure 11: Terracotta Lekythos, 550-530 BCE, Attic

the



Figure 13: Terracotta Lekythos, 550-530 BCE, Attic

individuals are also holding objects. The person on the left is holding a thin wand and the other is holding a feather-shaped object. With this information, one can infer that the square-shaped object behind these figures is a loom. The three in the middle are doing another job. The first one on the left is holding a narrow stick connected to four round objects. The next person appears as if they are either placing or taking away one of the round objects from the stick and into a trapezoid below. The person to their right has their mouth open and arms out. One can deduce that these figures are weighing the string which is tied up. The next two people seem like they are doing the same job as the second person from the left hand side, which is stretching out the wool into string. The last two figures are holding a rectangular cloth that they look like they are folding. Below the cloth, there are two other objects that look like already folded clothes sitting on a table. After looking at all of these scenes together, one can clearly see that it shows the process of making cloth and string. The first step is stretching the wool, then weighing it. Next, they weave it into cloth and, lastly, fold it. All of the figures in the horizontal, middle area of the lekythos are most likely women because of their styled hair and long dresses.

Using background knowledge, it is known that weaving was a female dominated hobby and job (Hemingway). The person that this funerary object was made for might have been an avid weaver, which explains why the lekythos is focused on the stages of making cloth. One can guess that all of the characters in this lekythos are women, as the clothing highly resembles a women's tunic. During this time, women would usually wear a floor-length tunic with a strophion, a waist band ("Ancient Greek Dress"). This funerary lekythos depicted scenes of a woman in power in the upper area and of women making cloth in the middle area. This lekythos was most likely placed in a woman's grave, considering how it primarily shows female characters. Usually, items that are placed in someone's grave carry a special meaning to them, so



Figure 14: Stela, 580-570 BCE, Laconian

these images probably describe a positive attribute of the person that the lekythos was made for. Bearing this in mind, the fact that it showed the process of weaving for funerary purposes shows that weaving was a very respected activity in Athens. Unlike in Sparta, where weaving would have been a job for helots, Athenian women spent much of their time in the house.

Figure 14 is a limestone stela from 580 to 570 BC that was found at the shrine of Helen and Menelaos (Elder). The relief is significantly damaged. The person on the left is supposed to be Menelaus, who is wearing a helmet and holding a sword pointed at Helen. One of his legs is positioned outwards to the right and he is laying his left arm on Helen's shoulder. Helen, who is on the right, is wearing a long veil that almost reaches the ground. Her hair is straight and down her back, which shows that she is still a maiden ("Stela"). She is holding a thin, round object resembling a wreath and handing it to Menelaus.

It is very much possible that this relief was meant to depict a story about the marriage between Helen and Menelaus. In this story, Helen goes off to Egypt, and, when Menelaus finally arrives, he sees Helen about to wed the Egyptian ruler. Because of this, Menelaus kills Deiphobus and almost slays Helen. However, he decides to forgive her and take her back as his wife (Gill). This narrative of Helen and Menelaus indirectly describes the convoluted relationship between them. Helen was once a goddess, but she lost her status and kept her feminine power (Nagy). Somehow, Helen, someone who no longer has her position as a goddess, is able to avoid Menelaus, a commanding king, to forgive her actions over and over again. Although she does not officially have a large position, her ability to continuously sway Menelaus is a reflection of her power which is overlooked. In this limestone stela, Menelaus and Helen are depicted as similar sizes, representing how they have the same amount of influence. Helen is standing up straight and not backing down, even though Menelaus has a sword, telling the viewer about her character. She is not afraid of her husband and is just as strong as him. Nancy Joseph writes, "In the case of Helen, the early feminist movement made Helen an unpopular character, with her power deriving from her beauty and femininity. But, says Blondell, feminism has more recently turned its attention to 'feminine' modes of power, such as the power of beauty and sexuality (Joseph)." Helen was an early symbol of feminism and has stayed as one, which she was heavily criticized for. Even today, many adapters make her seem like the antagonist in the story. Although Menelaus is furious about Helen's actions and feels the need to end her life, Helen is able to stop him.

Engaging with Alternate Perspectives

Many may argue opposing views of Spartan and Athenian women that contradict the primary thesis of this research paper, some of which will be provided in this section. One could say that the women in Sparta still did not have many rights compared to other women of that time. This could be reasoned as true when compared to women in other places at this time such as Egypt. However, for the purpose of this paper, the women of Sparta are being compared with the women in the region around them, specifically those of Athens. When put in this context, Spartan women did seem to have an incredible amount of power and were seen as equals to men.

One could also state that Athens and Sparta are similar in many ways. In this case, since Sparta and Athens are fairly close in distance, they will be somewhat alike no matter what, as shown through their art styles. They both use common Greek art techniques like decorated bronze mirror handles. Considering their proximity, though, the two are significantly divergent in most aspects. For instance, the two cities were governed extremely differently: Sparta was considered an oligarchy and Athens was a democracy.

Conclusion

In this research paper, I analyzed artworks relating to both Athenian and Spartan women to compare how they were treated. Unlike Spartan women, Athenian women were viewed as second class citizens with many less rights than men. Art was chosen for all of the sources because art has a large role in history. Art has been around for thousands of years, so there are many primary sources of art dating back to when Sparta existed. Most of the pieces come from specifically the fifth and sixth century BCE, since this was a peak artistic era. Sparta also had a limited amount of written material, so a majority of the writing that still exists comes from Athens. Therefore, using art actually from Sparta removes some bias and misinformation. Additionally, art can be interpreted in a variety of ways depending on the viewer and background knowledge. Learning about Sparta and Athens provides an insight into two contrasting belief systems of two city states that were relatively close in distance. One can understand their mindsets and actions and compare their thoughts to ours today. By doing this, one realizes how our beliefs evolve with time and circumstances. In today's world, opinions heavily differ from those of Athens and Sparta, but one can still learn from them in other ways. Society could learn from Athens by strongly valuing education and wisdom, while also learning from Sparta by caring for the physical body.

Works Cited

- Agnew, Molly Elizabeth. "Ancient Greek Beauty Standards." *Eternal Goddess*, 19 June 2022, www.eternalgoddess.co.uk/posts/ancient-greek-beauty-standards#:~:text=Women%20looked%20up%20to%20Aphrodite,and%20a%20pear%2Dshaped%20body. Accessed 19 Aug. 2024.
- "Ancient Greek Girl Athlete [Sculpture]." *Children and Youth in History*, chnm.gmu.edu/cyh/primary-sources/383.html#:~:text=This%20unusual%20bronze%20figurine%20of,public%20schooling%20including%20physical%20education. Accessed 6 Sept. 2024.
- "Artemis and Aphrodite." *Duke University*, [people.duke.edu/~wj25/Greek_Civ_docs/Artemis&Aphrodite.html#:~:text=Artemis%20%26%20Aphrodit&text=Artemis%20\(%2D%20Roman%20Diana\),hunt%2C%20chastity%2C%20and%20childbirth%20\(!\)&text=Associated%20with%20%22virgin%22%20nature%2C,chastity%20\(but%20also%20childbirth!\)](http://people.duke.edu/~wj25/Greek_Civ_docs/Artemis&Aphrodite.html#:~:text=Artemis%20%26%20Aphrodit&text=Artemis%20(%2D%20Roman%20Diana),hunt%2C%20chastity%2C%20and%20childbirth%20(!)&text=Associated%20with%20%22virgin%22%20nature%2C,chastity%20(but%20also%20childbirth!)). Accessed 26 Aug. 2024.
- Bonner, Robert J. "The Book Divisions of Thucydides." *Classical Philology*, vol. 15, no. 1, Jan. 1920, pp. 73-82. JSTOR, www.jstor.org/stable/263546.
- Christesen, P. *Athletics and Social Order in Sparta in the Classical Period*. JSTOR, <https://doi.org/10.1525/ca.2012.31.2.193>.
- Cole, Susan Guettel. *Could Greek Women Read and Write?* <https://doi.org/10.1080/00497878.1981.9978534>.
- Cosette. "Epithets of Athena, the Powerful Bright-Eyed Greek Goddess." *Divine Hours*, 22 July 2023, cosettepanequ.com/epithets-of-athena/.
- Day, Peggy L. "Anat: Ugarit's 'Mistress of Animals.'" *The University of Chicago Press Journals*, vol. 51, no. 3, July 1992, p. 181, www.journals.uchicago.edu/doi/abs/10.1086/373550?journalCode=jnes. Accessed 19 Aug. 2024.
- Denis, Paul. "A Greek Wedding." *ROM*, 29 Sept. 2019, www.rom.on.ca/en/collections-research/magazine/a-greek-wedding. Accessed 19 Aug. 2024.
- Department of Greek and Roman Art. "Ancient Greek Dress." *The Met*, Oct. 2003, www.metmuseum.org/toah/hd/grdr/hd_grdr.htm. Accessed 6 Sept. 2024.
- Elder, Ann. "The Shrine of Helen and Menelaos." *The Athenian*, 1 May 1993, www.the-athenian.com/site/1993/05/01/ther-shrine-of-helen-and-menelaos/. Accessed 19 Aug. 2024.
- "Eros/Amors/Cupid." *Ruks Museum*, www.rijksmuseum.nl/en/rijksstudio/subjects/erosamor/cupid#:~:text=Eros%20was%20the%20Greek%20god,an%20arrow%20into%20their%20heart. Accessed 7 Sept. 2024.
- Farnell, Lewis R. "Plato Comicvs: Frag. Phaon II.: A Parody of Attic Ritual." *The Classical Quarterly* 14.3-4 (1920): 139-146. Web.
- Figure*. 520-500 BCE, www.britishmuseum.org/collection/object/G_1876-0510-1?_gl=1%2af2egae%2a_up%2aMQ..%2a_ga%2aNTc1MTE1NTEJLjE3MjMzMTgwOTQ.%2a_ga_08TLB9R8X1%2aMTcyMzMxODA5NC4xLjAuMTcyMzMxODA5NC4wLjAuMA.
- Gill, N. S. "About the Trojan Prince Deiphobus." *ThoughtCo*, 20 May 2018, www.thoughtco.com/who-is-deiphobus-117980. Accessed 5 Sept. 2024.

- Greek and Roman Mythology Dictionary. "Corinna." *UPENN*,
www2.classics.upenn.edu/myth/php/tools/dictionary.php?method=did&regexp=1300&setcard=0&link=0&media=0. Accessed 19 Aug. 2024.
- "The Government of Ancient Athens." *EdBooks*,
edtechbooks.org/democracy/athens#:~:text=Only%20free%20adult%20men%20who,political%20rights%20or%20political%20power.
- Hemingway, Colette. "Women in Classical Greece." *The Met*, Oct. 2004,
www.metmuseum.org/toah/hd/wmna/hd_wmna.htm#:~:text=The%20gynaikon%20was%20where%20mothers,had%20some%20mobility%20in%20antiquity. Accessed 3 Sept. 2024.
- Herrington, C. J. "Athena in Athenian Literature and Cult." *Greece and Rome. JSTOR*,
www.jstor.org/stable/826896?searchText=athena&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3Dathena%26so%3Drel%26efqs%3DeyJjdHkiOlsiYW05MWNtNWWhiQT09IiwWTJoaGNIUmxjZz09IiwY21WelpXRnlZMmhmY21Wd2IzSjAiXX0%253D&ab_segments=0%2Fbasic_search_gsv%2Fcontrol&refreqid=fastly-default%3A114a50bb424e18f5555a653e92bb59fa.
- "History of Tapestries from Medieval to Modern Times." *Touch of Tapestry*,
touchoftapestry.com/History-of-Tapestries-7.html?srsId=AfmBOoqcV19D5mhgkaXAcQV3nFte8GF7hxWACvjswG6htiak_H9JhAOp. Accessed 7 Sept. 2024.
- Howitt-Marshall, Duncan. "What Was It like Being a Woman in Ancient Greece?" *Greece Is*, 2 Dec. 2022,
www.greece-is.com/like-woman-ancient-greece/#:~:text=Athenian%20girls%20were%20educated%20at,how%20to%20read%20and%20write.
- Joseph, Nancy. "Revisiting Helen of Troy." *College of Arts and Sciences*, 1 Sept. 2009,
artsci.washington.edu/news/2009-09/revisiting-helen-troy. Accessed 19 Aug. 2024.
- Kapparis, K. "Women and Family in Athenian Law." *STOA*, 22 Mar. 2003,
www.stoa.org/demos/article_women_and_family@page=all&greekEncoding=UnicodeC.html. Accessed 5 Sept. 2024.
- LaRocca-Pitts, Beth. "Anath: Bible." Jewish Women's Archive, 24 Aug. 2024,
[jwa.org/encyclopedia/article/anath-bible#:~:text=Anath%20\(Anat\)%20is%20a%20prominent,fertility%20and%20storm%20god%20Baal](http://jwa.org/encyclopedia/article/anath-bible#:~:text=Anath%20(Anat)%20is%20a%20prominent,fertility%20and%20storm%20god%20Baal). Accessed 26 Aug. 2024.
- "Let's Travel Back in Time: A Look at the Sanctuary through the Ages..." Women in Antiquity, 26 Nov. 2018, womeninantiquity.wordpress.com/2018/11/26/the-sanctuary-of-artemis-orthia/. Accessed 19 Aug. 2024.
- McKeown, Heather, and Debby Sneed. "Athenian Red-Figure Vase Painting." University of Colorado Boulder, 15 June 2018,
www.colorado.edu/classics/2018/06/15/athenian-red-figure-vase-painting#:~:text=This%20technique%20of%20vase%20painting,color%20of%20the%20base%20clay. Accessed 21 Aug. 2024.
- "Medias Painter." Getty Museum, www.getty.edu/art/collection/person/103JQZ. Accessed 19 Aug. 2024.
- Miller, Madeline. "Myth of the Week: Athena." Madeline Miller, 27 Feb. 2012,
madelinemiller.com/myth-of-the-week-athena/.

Mirror. 490-460 BCE, www.britishmuseum.org/collection/object/G_1854-0810-1.

"Mirrors." The Fitzwilliam Museum, fitzmuseum.cam.ac.uk/explore-our-collection/highlights/context/sign-and-symbols/mirrors. Accessed 24 Aug. 2024.

Nagy, Gregory. "Helen of Sparta and Her Very Own Eidolon." *Classical Inquiries*, 2 May 2016, classical-inquiries.chs.harvard.edu/helen-of-sparta-and-her-very-own-eidolon/. Accessed 5 Sept. 2024.

"Parthenon. East Frieze. Block 8 (VIII). Fragments." *Acropolis Museum*, www.theacropolismuseum.gr/en/parthenon-east-frieze-block-8-viii-fragments. Accessed 19 Aug. 2024.

"Peloponnesian War." *Britannica*, www.britannica.com/event/Peloponnesian-War.

The British Museum. "Temple-relief: Series: The Parthenon Sculptures," www.britishmuseum.org/collection/object/G_1816-0610-24.

Plutarch. "Lycurgus." Translated by John Dryden. *Knarf*, knarf.english.upenn.edu/Plutarch/Lycurgus.html. Accessed 7 Sept. 2024.

Reconstruction Drawing of Cult Statue of Athena Parthenos in Cella of Parthenon. 447-39 BCE. JSTOR, www.jstor.org/stable/community.13563252?searchText=&searchUri=&ab_segments=&searchKey=&refreqid=fastly-default%3Abab39367cd8ffd384cc0a6716317b0bc.

"Sanctuary of Athena Chalkioikos." *Lakonia Mobi*, lakonia.mobi/en/article/mnimeia/arxaioi-xronoi/iero-tis-athinas-xalkioikou/1709.

Sparta. Vol. 25, 1911, en.wikisource.org/wiki/1911_Encyclop%C3%A6dia_Britannica/Sparta.

"Sparta." *Lumen*, courses.lumenlearning.com/atd-herkimer-westerncivilization/chapter/sparta/#:~:text=Unique%20in%20ancient%20Greece%20for,rest%20of%20the%20classical%20world.

Spears, Betty. "A Perspective of the History of Women's Sport in Ancient Greece." *Journal of Sport History*. JSTOR, <https://www.jstor.org/stable/43609020>.

Stela. 330-317 BCE, www.britishmuseum.org/collection/object/G_1909-0611-1.

Stela. 580-570 BCE. *ARTstor*, www.jstor.org/stable/community.11657540?searchText=&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3Dsparta%26image_search_referrer%3Dhomepage%26so%3Drel%26searchkey%3D1721155273656%26doi%3D10.2307%2Fcommunity.11657541&ab_segments=0%2Fbasic_image_search%2Fcontrol&searchKey=&refreqid=fastly-default%3Af0e707e1a8494710ed7f9f407297602e.

Stewart, Andrew F. "Phidias." *Oxford Classical Dictionary*. December 22, 2015. Oxford University Press. Date of access 18 Nov. 2024, <https://oxfordre.com/classics/view/10.1093/acrefore/9780199381135.001.0001/acrefore-9780199381135-e-4954>

The British Museum. "Bronze Figure of a Running Girl, Sparta, 6th C. BC." *The British Museum*, www.bmimages.com/preview.asp?image=00075440001. Accessed 19 Aug. 2024.

The Metropolitan Museum of Art. "Bronze Mirror with a Support in the Form of a Nude Girl. 6th Century BCE," www.metmuseum.org/art/collection/search/253556. Accessed 5, Sept. 2024

The Metropolitan Museum of Art.. "Lead Figure of a Winged Goddess. 7th-6th Century BCE," www.metmuseum.org/art/collection/search/251570.

The Metropolitan Museum of Art. "Terracotta Lekythos (oil flask), 550-530 BCE," www.metmuseum.org/art/collection/search/253348. Accessed 5, Sept. 2024

Thompson, M. S. "The Asiatic or Winged Artemis." *The Journal of Hellenic Studies*. [JSTOR], www.jstor.org/stable/624687?oauth_data=eyJlbWFpbCI6InZpdmlbm5lLnRhbmcuMjdAc2FjcmVkc2Yub3JnIiwiaW5zdGl0dXRpb25JZHMlOltLCJwcm92aWRlcil6Imdvb2dsZSJ9.

Thucydides. *History of the Peloponnesian War*. United Kingdom, Edwards & Bushnell, 1856.

Tod, Marcus Niebuhr. *1911Encyclopædia Britannica*, https://en.wikisource.org/wiki/1911_Encyclop%C3%A6dia_Britannica.

Tomlinson, Richard Allan. *Parthenon*. 2016, [oxfordre-com.yale.idm.oclc.org/classics/view/10.1093/acrefore/97801\[...\]acrefore-9780199381135-e-4758](http://oxfordre-com.yale.idm.oclc.org/classics/view/10.1093/acrefore/97801[...]acrefore-9780199381135-e-4758).

"The Two Faces of Greece: Athens and Sparta." *PBS*, www.pbs.org/empires/thegreeks/educational/lesson1.html#:~:text=Spartan%20culture%3A%20Militaristic%20values,elderly%2C%20women%2C%20and%20warriors. Accessed 19 Aug. 2024.

Vanderpool, Eugene, and Blake Ehrlich. "Athens." *Britannica*, 26 Sept. 2024, www.britannica.com/place/ancient-Greece/Sparta-and-Athens.

"Vase." *EVL*, www.evl.uic.edu/animagina/symbols.html#:~:text=vase&text=It%20also%20stands%20as%20the,is%20also%20a%20maternal%20symbol. Accessed 7 Sept. 2024.

Vessel (Lebes Gamikos or Nuptial Lebes; Red-Figure with Additional Plastic and Gilded Decoration; Ht. 18 1/16"). Terra cotta., ca. 350 B.C. Leningrad: Mus., Hermitage; No. 15592.; Found on Thasos island., *JSTOR*, <https://jstor.org/stable/community.11658068>. Accessed 18 Nov. 2024.

Vessel (Hydria; Red-Figure Enriched with Gilt; Ht. 18 1/2"; Diam. of Mouth 6 1/2"). Terra cotta (detail)., ca. 410 B.C. Florence: Mus. Archeologico; No. 81947., *JSTOR*, <https://jstor.org/stable/community.11657830>. Accessed 18 Nov. 2024.

Vessel (Hydria; Red-Figure Enriched with Gilt; Ht. 18 1/2"; Diam. of Mouth 6 1/2"). Terra cotta (detail)., ca. 410 B.C. Florence: Mus. Archeologico; No. 81947., *JSTOR*, <https://jstor.org/stable/community.11657829>. Accessed 18 Nov. 2024.

Vessel (Hydria; Red-Figure Enriched with Gilt; Ht. 18 1/2"; Diam. of Mouth 6 1/2"). Terra cotta (detail)., ca. 410 B.C. Florence: Mus. Archeologico; No. 81947., *JSTOR*, <https://jstor.org/stable/community.11657828>. Accessed 18 Nov. 2024.

"What Was Most Valuable to Athenians?" *KNILT*, 9 Sept. 2010, knilt.arcc.albany.edu/Lesson_2:_What_was_most_valuable_to_Athenians%3F#:~:text=The%20Goddess%20Athena-,Athens%20Values,Boys%20were%20educated. Accessed 14 Aug. 2024.

Willette, Dorothy. "The Enduring Symbolism of Doves." *Biblical Archaeology Society*, 16 July 2024, www.biblicalarchaeology.org/daily/ancient-cultures/daily-life-and-practice/the-enduring-symbolism-of-doves/. Accessed 23 Aug. 2024.

"Worshipping Aphrodite." *The J. Paul Getty Museum*, www.getty.edu/art/exhibitions/aphrodite/worship.html.

"Xenophon." *Great Thinkers*, thegreatthinkers.org/xenophon/biography/.

Is Chemosynthetic Technology Viable and Realistic to Use in the Context of How Much Energy It Can Produce and its Environmental Impact? By Harika Chodavarapu

Abstract

Chemosynthesis is a process used by certain organisms at the bottom of the ocean near hydrothermal vents and cold seeps to convert chemicals into sugars used for energy. One type of chemosynthesis converts methane and oxygen into water, carbon dioxide, and glucose, which may be desirable to reduce the impact of methane as a greenhouse gas on the environment. This project aims to design biomimetic technology inspired by this biological process. This project focuses mostly on the viability and sustainability of the technology and presents a preliminary mechanical design for the device. The final process design has five reactions, ultimately producing a final product glycerol. While the design was not economically or spatially viable in its current form and thus did not meet all constraints, it is stoichiometrically possible, meeting two of three criteria set. Further research would include electricity usage reduction and common, cost-effective reactions used in industry.

Introduction

According to the UN Environment Programme, “methane traps more heat in the atmosphere per molecule than carbon dioxide (CO_2), making it 80 times more harmful than CO_2 for 20 years after it is released” (“What’s the Deal with Methane?”). Methane must not be released or allowed to stay in the atmosphere for its effect to be mitigated.

Chemosynthesis is a process akin to photosynthesis that happens without light. It converts chemicals into energy for organisms to use. Some microorganisms at the bottom of the ocean near hydrothermal vents or cold seeps use chemosynthesis to produce energy and sugar (“Chemosynthesis”). Methane-fueled chemosynthesis converts methane and oxygen into water, carbon dioxide, and glucose through a two-step process that involves methane combustion and carbon fixation. The methane and oxygen are first converted into methanol, carbon dioxide, water, and chemical energy. Then, the methanol is converted into formaldehyde, which undergoes a carbon fixation process that uses the chemical energy produced by the first part of the process to become glucose (Figure 1) (“Chemosynthesis”). Chemosynthesis is not currently being mimicked by human technology to produce energy or other useful products, so this project aims to mimic chemosynthesis in order to both devolatilize methane and produce glycerol in a sustainable fashion.

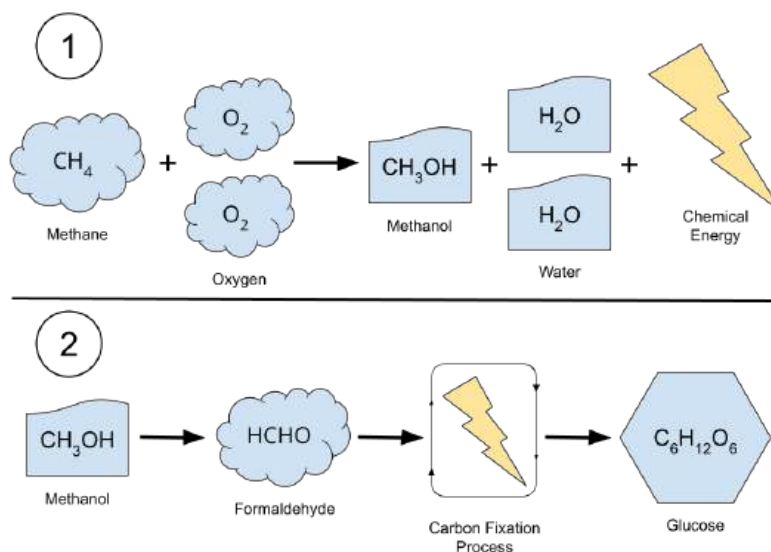


Figure 1: Methane-based chemosynthesis

It may be desirable to produce glycerol, or glycerin, which can be used in many ways, such as to make soaps, fuels, and as a pharmaceutical precursor (“Glycerol”). Glycerol could be produced sustainably by placing the hypothetical technology on a cow farm because of its abundance in methane. The United States Environmental Protection Agency (EPA) states, “A single cow produces between 154 to 264 pounds of methane gas per year. Not including the emissions of any other livestock, 1.5 billion cattle, raised specifically for meat production worldwide, emit at least 231 billion pounds of methane into the atmosphere each year (Our World in Data)” (“Agriculture and Aquaculture”). Having the technology on a cow farm presents the opportunity to make the system circular in that the cows produce methane, which fuels the technology to make glycerol, which could potentially be used for the cows or could be sold for the farm’s profit. This could enable reduction of agricultural methane emissions.

The four constraints for the design of the technology were that the theoretical volume of the device should not exceed 40 m³, the calculated cost should not exceed \$100,000 to make the device, the calculated cost of energy for this device should not exceed \$10,000 per year, and each cycle of the hypothetical device should not take more than a day. The volume constraint was made to consider the available space on a farm. The manufacturing cost constraint and the energy cost constraint were made to consider the budget available to a farm. The time constraint was made so that there can be some output from the device per day. To assess how well the mechanical design works, there were three success criteria: the hypothetical device should have a calculated input of at least 250 kg of methane in one day, the hypothetical device should have a calculated output of at least 100 kg of glycerol in one day, and the calculated monetary value of the glycerol produced per year should exceed the calculated cost of the energy used per year.

Overall, this project aims to assess the viability and sustainability of a chemosynthesis-inspired technology and create a rudimentary mechanical design for the technology. This work is a theoretical exploration into the feasibility of such a chemosynthesis-based technology that would convert methane into glycerol and does not build any of the technology discussed. To achieve these goals, the technology would be based on methane-fueled chemosynthesis and would hypothetically be placed in a cow farm.

Materials and Methods

The final conceptual design had five reactions. The first reaction would turn methane into syngas (Alvarez-Galvan, Consuelo, et al.). The second would turn the syngas into methanol (Timsina, Ramesh, et al.). The third would turn the methanol into formaldehyde (Thrane, Joachim, et al.). The fourth would turn the formaldehyde into dihydroxyacetone (DHA) (Deng, Jin, et al.). The fifth and final reaction would turn the DHA into glycerol (Gracey, Benjamin Patrick, et al.).

The general process to design the theoretical device started with figuring out the mass balances, or stoichiometry, for all of the reactions. Balancing the masses means making sure that all the mass that entered a reaction is accounted for in the products of the reactions, following the law of conservation of mass. For example, the chemical equation of the first part of chemosynthesis is: $\text{CH}_4 + 2 \text{O}_2 \rightarrow \text{CH}_3\text{OH} + 2 \text{H}_2\text{O}$. This means that if one mole of methane enters the reaction, two moles of oxygen also have to enter, and one mole of methanol and two moles of water have to exit the reaction. This is because one carbon atom, four hydrogen atoms, and four oxygen atoms enter the reaction, so one carbon atom, four hydrogen atoms, and four oxygen atoms have to exit the reaction to maintain mass conservation. After the mass balances were found for all reactions, the energy required to fuel the reactions was found, followed by the power required. Finally, the volume and hardware were found, and the final values were used to evaluate the design according to the constraints and criteria. When finding the cost of the device, the costs of reactants, products, and catalysts were not evaluated.

The first two criteria for the design's success were that the hypothetical device should have a calculated input of at least 250 kg of methane in one and a calculated output of at least 100 kg of glycerol in one day. First, it was assumed that the input of the entire device was 10 kg (623.3 mol) of methane in order to scale the reaction down to make the values easier to calculate. Then, using the scales provided by the resource used for each reaction, the number of moles and the number of kilograms of each reactant, product, and catalyst there is for each reaction was calculated. The selectivities for each reaction were found in the literature and used to find more accurate amounts of useful products formed from the reactants of each reaction. It was also assumed that no external matter was being input into or output from the reactions and that only the mentioned reactants and catalysts formed the mentioned products. Then, the energy and power required for each reaction was calculated, as detailed later. After the mass values were found, each reactant, catalyst, and product amount was scaled up so that the amount of methane input was greater than or equal to 250 kg (15583.6 mol) and the amount of glycerol was greater

than or equal to 100 kg (1085.9 mol glycerol). Once the volumes of each reactant, catalyst, and product was calculated, as detailed later, it was determined if it would be reasonable to fit all of the reactants, catalysts, and products into the device under the volume constraints. If so, the concept design met both of these criteria. As an extra step for more information about the significance of the amount of glycerol produced, the value of the glycerol was calculated.

One of the constraints was that the calculated cost of power for the hypothetical device should not exceed \$10,000 per year. To evaluate this, after the scaled-down mass calculations were complete, the amounts of energy in kilojoules (kJ) necessary to heat the reactants and catalysts of each reaction to their respective required temperatures as stated in the sources were calculated. It was assumed that there was no additional energy needed and no energy lost. To do this, the specific heat of each reactant and catalyst was found and multiplied by the mass and then by the temperature change between the last reaction's temperature (or room temperature if it was the first reaction) and the temperature of the current reaction. If further information about the catalyst could not be found, no energy amount was calculated for temperature change, or the specific heat at a different starting temperature was used. After this energy was calculated, the amounts of energy required to pressurize the reactants and catalysts of each reaction to their respective required pressures as stated in the sources were calculated. To do this, the initial density of each reactant at the temperature and pressure of the previous reaction (or room temperature and normal pressure if it was the first reaction) was found. Then, the volume of each reactant and catalyst before pressurization was found using its initial density. The energy for pressurization was not found for most catalysts due to insufficient available information, so it was found only for the catalyst of the last reaction. The pressure of each reaction was found within its source article, and if it was not mentioned in the article, then the pressure of the reaction was assumed to be standard pressure. After, this pressure was used to find the final volume of each reactant and catalyst after pressurization. Then, the initial volume of each reactant was subtracted from its corresponding final volume, and this quantity was subsequently multiplied by the negative value of the pressure to receive the resulting energy required for pressurization. After the values of the energy required for the temperature change and the energy required for pressurization were found, these values were added together to receive the amount of energy required for both to occur. The energy values were scaled up to the full-size reaction by multiplying by a factor of 25.

After all the energy values were found, the power of each energy measurement was calculated based on the time it would take to heat or pressurize the reactants and catalysts, and this time was assumed to be the same as the time each reaction, or part of the reaction, took. Another constraint was that the time for one full cycle of the device could not take more than a day to complete, and all of the times for the reactions were added together for this constraint to be evaluated. It was assumed that there were no interruptions in the cycle, and the time for making the catalysts was assumed to be irrelevant. Since the amount of time the second reaction took was not listed in its literature, it was assumed to take one hour, and since the time the third reaction took was also not explicitly stated in its literature but there was a figure displaying

catalyst selectivity over a large amount of time, the third reaction was assumed to take five hours. Once the power was found in kilowatt-hours (kWh), it was scaled up to the full-size reaction by multiplying by a factor of 25. Then, the total power for one cycle of the device was multiplied by 365 to account for the power necessary to run the device for one cycle a day for a whole year. This value was multiplied by the standard rate for power in the United States, which was \$0.18 per kWh (“Electricity Cost in California”). If this cost was less than or equal to \$10,000, the concept design fit the design constraint.

Another design constraint evaluated was for the calculated volume to not exceed 40 m³. The volume of each reactant, catalyst, and product was calculated at its highest temperature and lowest pressure described in previous calculations using the scaled-down values at first. If the density or volume could not be found for a certain reactant, catalyst, or product at its highest temperature and/or its lowest pressure but its density or volume could be found at a different temperature and/or pressure, that density or volume was used for the final calculations. If no density or volume could be found, then the reactant, catalyst, or product’s volume was not evaluated in the final count. Once the rest of the volumes were found, they were scaled up by a factor of 25 and added up to find a final estimate of the volume without considering the volume of the equipment required. If this final estimate was less than or equal to 40 m³, then this constraint was met.

Finally, the last criterion the design needed to meet to be successful was that the calculated monetary value of the glycerol produced per year should exceed the calculated cost of the energy used per year. This was to ensure profitability for the user. To assess this, the mass of the glycerol produced by the device was multiplied by the number of cycles per year and then multiplied by \$1.13/kg of glycerol (“Glycerol Price Index”). If this price exceeded the cost of power calculated previously, then the criterion was met. An assumption was made that no mass would be lost in any purification the glycerol may go through after exiting the device.

Additionally, all values were calculated and recorded in a spreadsheet (see Supporting Information).

Results

Some assumptions were made during the calculations when sufficient information could not be found about the density at the highest temperature of the reaction or something similar, as listed in the Materials and Methods section. The final process design is summarized below in Figure 2 and Table 1. This multi-process reactor design did follow the two success criteria about at least 250 kilograms of methane as input and at least 100 kilograms of glycerol as output. It also met the constraint of each cycle happening in a day or less. However, it did not meet the cost criterion or the cost or volume constraints. The total cost of power for this device would be around \$225,652.15 per year using the average cost of power in the U.S. in early 2024. The total monetary value of the glycerol would be around \$74,504.24 per year. The total volume of the device including its hardware would be around 306.42 cubic meters, rounded up. Sufficient information could not be found about manufacturing costs to qualify the constraint.



Figure 2: Flowchart of final conceptual design reaction values

Table 1: Final conceptual design total values

Total Hardware Volume (m ³ , rounded down):	310
Energy per Cycle (kJ, rounded up):	12,364,502
Power per Cycle (kWh, rounded up):	3,435
Power Cost per Year (\$0.18/kWh, rounded down):	\$225,652
Monetary Value of Glycerol Produced per Year (\$1.13/kg, rounded down):	\$74,504
Net CO ₂ Removal per Year (kg, rounded up):	132,067
CO Output per Year (kg, rounded down):	46
H ₂ O Output per Year (kg, rounded down):	305

Discussion

Based on the process design results shown in the previous section, I determined that a chemosynthesis-inspired device is indeed possible to produce physically. However, the process is not viable due to money, power, and space constraints. The calculated cost of power (\$226,000) was much greater than \$10,000 per year, especially because of the first, second, and third reactions' power needs. In addition, the volume constraint was not met, primarily due to the volume requirements of the formaldehyde in the fourth reaction (267.55 m³), exceeding the constraint of 40 m³. The time constraint was met, however, because the calculated time each cycle of the device would take was less than a day (≈ 16 hours). Sufficient information could not be found about the necessary equipment to calculate the manufacturing cost of the device. In addition, the criteria were not all met either. Although the device's input and output standards were both met stoichiometrically, the criterion that stated that the monetary value of the yearly end product must exceed the energy cost per year was not met due to the steep cost of the energy necessary per year and the less valuable glycerol end product. Thus, the constraints and the criteria were not met according to the mathematics used to estimate the converter's effectiveness.

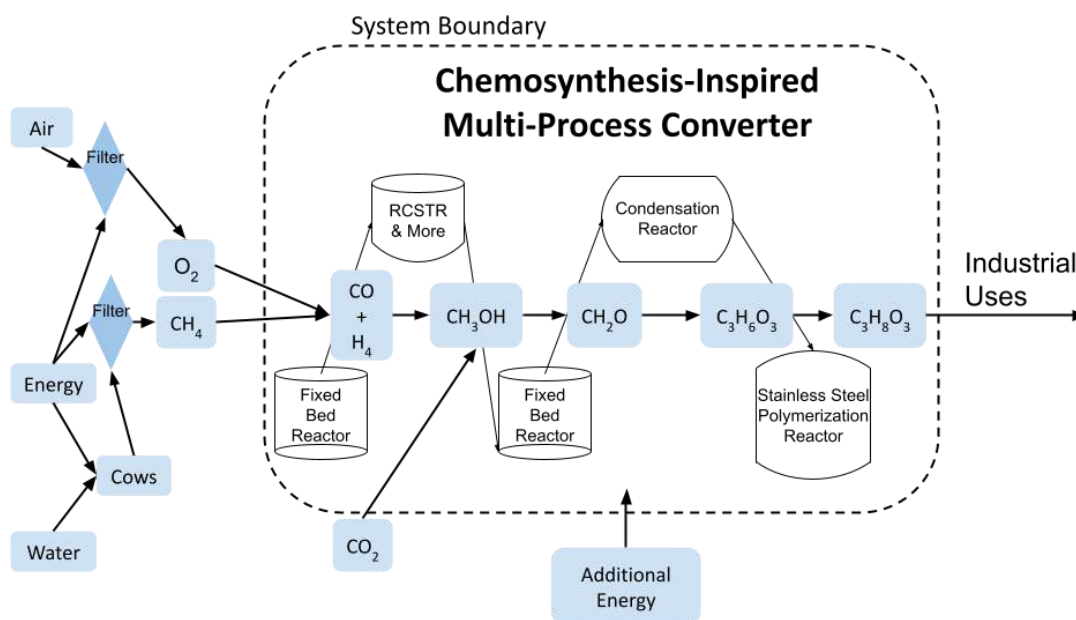


Figure 4: Flowchart of basic final conceptual design system

To improve this design, more industrially-commonplace reactions could be more suitable to solve the issues of volume and energy cost, especially for the energy-intensive second reaction that turns syngas into methanol and the volume-intensive fourth reaction that turns formaldehyde into DHA, because common industrial reactions tend to be optimized for efficiency. The use cases of the device may also be reconsidered to focus on something at a smaller scale than a cow farm, and the device should be scaled accordingly. Alternative pathways for the formaldehyde to turn into other products using reactions that are less taxing should also be considered.

Works Cited

- “What’s the Deal with Methane?” UNEP, UN Environment Programme,
www.unep.org/news-and-stories/video/whats-deal-methane.
- “Chemosynthesis.” NOAA Ocean Exploration,
oceanexplorer.noaa.gov/edu/materials/chemosynthesis-fact-sheet.pdf.
- “Glycerol.” Glycerol - an Overview | ScienceDirect Topics,
www.sciencedirect.com/topics/agricultural-and-biological-sciences/glycerol.
- “Agriculture and Aquaculture: Food for Thought.” EPA, Environmental Protection Agency, Oct. 2020, www.epa.gov/snep/agriculture-and-aquaculture-food-thought.
- Alvarez-Galvan, Consuelo, et al. “Partial oxidation of methane to syngas over nickel-based catalysts: Influence of support type, addition of rhodium, and preparation method.” *Frontiers in Chemistry*, vol. 7, 13 Mar. 2019, <https://doi.org/10.3389/fchem.2019.00104>.
- Timsina, Ramesh, et al. “Methanol synthesis from syngas: A process simulation.” *SIMS EUROSIM 2021*, vol. 185, Sept. 2021, pp. 444–449, <https://doi.org/10.3384/ecp21185444>.
- Thrane, Joachim, et al. “Highly stable apatite supported molybdenum oxide catalysts for selective oxidation of methanol to formaldehyde: Structure, activity and stability.” *ChemCatChem*, vol. 13, no. 23, 27 Oct. 2021, pp. 4954–4975, <https://doi.org/10.1002/cctc.202101220>.
- Deng, Jin, et al. “Linked strategy for the production of fuels via Formose reaction.” *Scientific Reports*, vol. 3, no. 1, 7 Feb. 2013, <https://doi.org/10.1038/srep01244>.
- Gracey, Benjamin Patrick, et al. *Synthesis of Glycerol from Formaldehyde*. 8 Mar. 1989.
- “Electricity Cost in California: 2024 Electric Rates.” EnergySage,
www.energysage.com/local-data/electricity-cost/ca/.
- “Glycerol Price Index.” Business Analytiq, 17 Sept. 2023,
businessanalytiq.com/procurementanalytics/index/glycerol-price-index/.

Predictive Modeling of Stock Market Prices Using Machine Learning By Aidan Lightman

Summary (Abstract)

The stock market plays a crucial role in the economy by allowing investors to buy and sell shares of publicly traded companies, reflecting various economic factors and serving as a basis for generating profits. Stock prices fluctuate based on company performance, market conditions, and external events, making it a dynamic and complex system. With the emergence of technologies like artificial intelligence (AI), there is growing interest in using AI to analyze and predict stock market movements. AI's ability to process large datasets, identify patterns, and make predictions has made it a valuable tool for investors seeking to improve decision-making and manage risks. By leveraging machine learning and neural networks, AI can provide more accurate forecasts and automate trading strategies.

In research, Python was used to build and evaluate several predictive models for stock prices, selecting the best-performing model based on mean squared error. This approach enabled the creation of an effective model for closely predicting stock price trends. After importing the necessary data, three types of predictive models were developed: a linear regression model, a neural network model, and a decision tree model. The linear regression model establishes a relationship between input and output variables by fitting a straight line through the data points. The neural network model simulates the way the human brain processes information through interconnected layers of nodes (neurons). The decision tree model splits data into branches based on specific questions and conditions.

Splitting the dataset into training and testing sets is a crucial step in machine learning to evaluate how well a model performs. By dividing the data, the model is trained on one part and tested on another, ensuring it can work well with new, unseen data. This approach prevents the model from memorizing patterns specific to one dataset (overfitting) and ensures it can generalize to other data in the future. Splitting the data also helps fine-tune the model's settings and measure its accuracy, making it easier to choose the best model. For each model, the data was split into training and testing sets, and the mean squared error (MSE) was used to measure accuracy.

After evaluating the performance of all three models, the best one was selected for further simulations based on its prediction accuracy. The linear regression model was determined to be the most effective. This model shows potential for profitability. For instance, an initial investment of \$1,000 in a single stock resulted in a return of \$2,677. The model was tested with various stock prices and quantities, consistently achieving positive outcomes.

Introduction

Stocks play a significant role in today's economy, reflecting every economic factor and change while providing a basis for investors to profit. The stock market is a platform where investors buy and sell shares of publicly traded companies. Purchasing a stock grants a small ownership stake in that company, and the stock's value fluctuates based on the company's

performance, investor demand, and market conditions. Investors make a profit by selling a stock at a higher price than it was purchased or incur a loss if sold for less.

Companies use the stock market to raise capital by selling shares to the public, which supports their expansion, innovation, and operational improvements. The stock market operates on exchanges like the New York Stock Exchange (NYSE) and NASDAQ, where buyers and sellers trade stocks at agreed-upon prices. Stock prices vary due to numerous factors, including company performance, market sentiment, economic indicators, and global events. For instance, strong earnings can increase a company's stock price, while concerns such as an economic recession can depress stock values across the board.

There are various strategies that investors employ in the stock market, ranging from short-term trading to long-term investing. Short-term trading is often based on market trends, while long-term investing focuses on company fundamentals. The stock market is seen as a barometer of economic health, reflecting investor confidence in both individual companies and the broader economy. However, it is also volatile, influenced by factors such as geopolitical tensions, technological advancements, and government policy changes. Despite these risks, the stock market offers substantial potential for wealth generation, requiring knowledge, research, and a balanced approach to navigate its uncertainties and fluctuations.

Emerging technologies are creating new opportunities to capitalize on financial markets. Artificial intelligence (AI) is at the forefront of this technological innovation, being used as a tool in many fields around the world. AI has been increasingly integrated into stock trading and investment strategies, analyzing vast amounts of historical data, identifying patterns, and making predictions that are not easily discernible to human traders. By leveraging techniques such as machine learning, natural language processing, and neural networks, AI can forecast stock prices, detect trends, and even automate trading decisions. This enables investors and financial institutions to make more informed decisions, manage risks better, and potentially increase returns.

Python, a versatile programming language commonly used for tasks like data analysis, web development, and automation, was utilized to build multiple predictive models for stock market prices. Various models were evaluated based on their mean squared errors to identify the most accurate predictions. The resulting model proved effective in predicting stock price trends for different companies, demonstrating the potential for AI to closely forecast market movements and inform trading strategies.

Materials and Methods

Various software tools were utilized for this project, which was executed on a web browser using a programming language called Python. Python is a versatile language commonly used for tasks like building websites, analyzing data, and automating processes. Its simplicity and flexibility make programming accessible for many developers. The project began by importing three essential Python libraries. The yfinance library provides easy access to historical stock market data from Yahoo Finance, making it convenient to download financial information

for analysis. Panda is a powerful data manipulation library used primarily to handle structured data such as tables or time-series data. NumPy is used for numerical operations and supports arrays, matrices, and mathematical functions.

The setup involved creating a Ticker object for Google's stock. This object allows for convenient access to historical and real-time stock data. The `history()` method was then called on the ticker object to retrieve Google's historical stock data over the last five years. This data was stored in a DataFrame. Unnecessary columns were dropped from the DataFrame, retaining only the 'Open' and 'Low' columns for further analysis. Next, two empty NumPy arrays, 'open' and 'close', each with a shape of 1259 elements (the approximate number of trading days over 5 years), were pre-allocated and set to float. However, these arrays were found to be redundant, as the Data Frame columns could be directly converted to NumPy arrays without pre-allocation. The 'Open' prices from the Data Frame were then converted to a NumPy array using the `.to_numpy()` method and assigned to the open variable for easier data manipulation in further analysis.

The next step involved opening the last three stock prices from Google to get a general sense of the values. Then, code was written to initialize two variables. After that, it created a list of 1260 zeros for labels or targets. The code iterated through a range of 1,250 elements, creating input and output datasets for a machine learning model. For each iteration, the code constructed an input vector consisting of three sequential opening stock prices. These values represented the opening prices for three consecutive trading days, capturing a short historical window of stock movements. The corresponding output value was set as the opening price on the fourth day, following the three-day sequence. This approach created a supervised learning dataset where each set of three consecutive days' opening prices predicted the opening price on the next day, stored. By sliding this window across the data, multiple training samples were generated, making it possible to train a model to predict future opening prices based on recent historical data.

Next, two essential libraries for data manipulation and machine learning were imported for numerical operations and for splitting the dataset into training and testing subsets. The datasets were then split accordingly, and the models were trained and tested.

The first model constructed was a linear regression model. The mean squared error function was also imported to evaluate the model's performance. A linear regression model was created and trained on the training using the `.fit()` method. After training, the model was used to make predictions on the test data, and the predicted values were stored. The model's coefficients were printed, and the mean squared error (MSE) was calculated between the actual test values and the predicted values. The MSE was displayed and stored in the linear MSE. Finally, the inverse MSE was calculated as the reciprocal of the mean squared error, which can be useful for comparing model performance—higher values of the inverse MSE indicate lower error and better model fit.

The second model built was a neural network. The model, named was initialized with a fixed random state for reproducibility and a maximum of 500 iterations. The `.fit()` method trained the model using the input features and the corresponding target values. After training, the model

predicted the target values for the test set, and the predictions were stored. The MSE between the actual test values and the predicted values was then calculated and printed, and inverse MSE of the neural network was computed as the reciprocal of the MSE.

The third and final model was a decision tree regressor. The model was created with a maximum depth of 2 to prevent overfitting. The model was trained on the training data and then used to make predictions on the test data, storing the results. The MSE was calculated between the actual test values and the predicted values and stored in MSE of the decision tree regressor. Lastly, inverse MSE of the decision tree regressor was computed as the reciprocal of the MSE, serving as an inverse measure of model error.

After building all three models, simulations were conducted using the model with the lowest mean squared error (MSE), which was the linear regression model. The simulation started by initializing two variables: amount, representing the initial cash available for trading, and stocks, indicating the number of stocks held. The trained linear regression model was used to predict future stock prices for the given input data. A basic trading strategy was simulated based on the model's predictions, iterating through 1,240 data points. At the end of the loop, the final cash value was calculated by selling any remaining stocks at the current price and adding this amount to the total cash value. The result provided an overview of the profit or loss achieved using this trading strategy.

To identify the best-performing model, a simple average of predictions from the three models was calculated. The mean squared error (MSE) of this averaged prediction was then computed for evaluation. Next, a weighted ensemble of the three models was created based on the inverse of their respective mean squared errors. Each model's predictions were weighted according to their performance, and the combined prediction, named combined, was calculated as a weighted sum of the three models' predictions. The mean squared error of the combined model was computed and compared against the individual models and the simple average model to identify the most accurate approach. We then simulated the trading bot. We would buy a certain number of stocks and it would sell them when a reasonable profit was made. Thus, the bot became successful using this strategy.

Discussion

Our primary objective in developing this model was to increase our ability to capitalize on opportunities in the stock market, thereby enhancing our chances of generating consistent and significant returns. By leveraging advanced data analysis and predictive algorithms, we aim to make more informed decisions, reduce risk, and ultimately achieve greater financial success.

The project started with the inclusion of three essential libraries that enabled the collection, manipulation, and processing of data. Historical stock market data for Google's stock over the past five years was collected and refined into a structured dataset, making it suitable for analysis.

Three different models were developed to analyze the data: a linear regression model, a neural network, and a decision tree regressor. The performance of each model was measured

using their prediction accuracy, and the linear regression model proved to be the most effective, making it the primary model for simulations. Using predictions from this model, a trading strategy was created that simulated stock purchases and sales based on predicted price movements, providing insights into potential profits or losses.

To further enhance the predictive capabilities, two ensemble models were developed. The first was a simple average of predictions from all three models, while the second used a weighted approach that prioritized models with better accuracy. This comprehensive strategy enabled a detailed evaluation of each model's effectiveness, ultimately identifying the best-performing model for predicting future stock prices. Using this approach, for example, when we start with \$1000 for a Google stock and get three stocks, we can make a \$1654 profit. Thus, this shows that you can make money from our model because of the profit made from each of our tests.

Results

MSE	Google	Tesla	Apple	Gold
Linear Regression Model	3.98	79.33	7.16	.19
Neural Network	5.42	79.30	7.42	.22
Decision Tree Regressor model	7.72	101.03	10.90	.24

This table represents the Mean Squared Error from each model for each different company.

1 Stock

Profit (\$)	Google	Tesla	Apple	Gold
1000	1559	18947	2665	1336
5000	7237	38391	11618	2765
10000	14509	55506	22998	3288

This table represents the profit gained for each company when you buy and sell one stock using my linear regression model.

3 Stocks

Profit (\$)	Google	Tesla	Apple	Gold
1000	1654	19597	2538	1961
5000	7958	69753	12888	5537
10000	15320	88265	24886	6757

This table represents the profit gained for each company when you buy and sell three stocks using my linear regression model.

Additional Photos of Code

```
import matplotlib.pyplot as plt
import numpy as np
from sklearn import datasets, linear_model
from sklearn.metrics import mean_squared_error, r2_score
Linear = linear_model.LinearRegression()
Linear.fit(X_train, Y_train)
Y_pred = Linear.predict(X_test)
print("Coefficients: \n", Linear.coef_)
print("Mean squared error: %.2f" % mean_squared_error(Y_test, Y_pred))
mse_linear = mean_squared_error(Y_test, Y_pred)
mse_linear_inverse = 1 / mean_squared_error(Y_test, Y_pred)
```

This picture represents the code for setting up the linear regression model.

```
import yfinance as yf
import pandas as pd
import numpy as np

#setup
#code
ticker = yf.Ticker('GOLD') #import data
aapl_df = ticker.history(period="5y") #get data from 5 year period in dataframe
aapl_df.drop(['High', 'Close', 'Volume', 'Dividends', 'Stock Splits'], axis=1, inplace=True)
open = np.empty(shape = (1259), dtype = float)
close = np.empty(shape = (1259), dtype = float)

open = aapl_df[['Open']].to_numpy()
```

This picture represents the code for importing the datasets and setting up the basis for the project.

```
amount=1000
stocks=0
prediction=Linear.predict(X)
for i in range(1240):
    if Y[i] < prediction[i+1] and amount > Y[i]:
        amount=amount-Y[i]
        stocks=stocks+1
    elif Y[i] > prediction[i+1] and stocks > 0:
        amount=amount+Y[i]
        stocks=stocks-1
    if i==1239:
        amount=amount+stocks*Y[i]
print(amount)
```

This photo is the code for a trading simulation that we used to gain profit. IN this particular instance, this code buys one stock and sells one stock.

Works Cited

Stock Price Predictions With a Linear Model and Neural ...Journal of Student

Research<https://www.jsr.org> › index.php › path › article › view

Stock market prediction using artificial

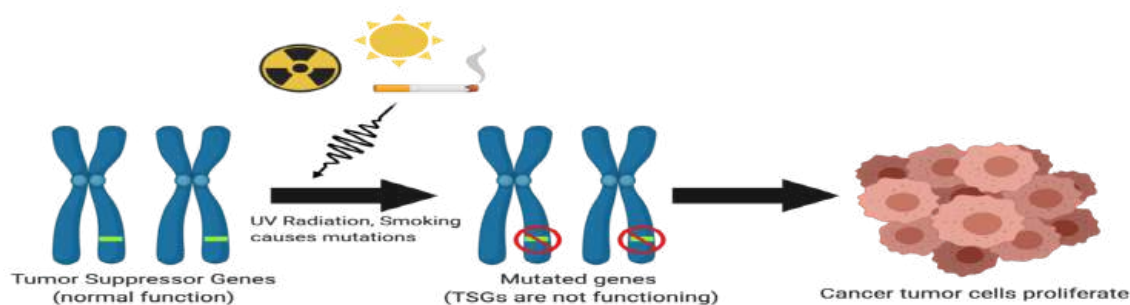
intelligenceScienceDirect.com<https://www.sciencedirect.com> › science › article › pii

The Cardiovascular Implications of Mutated Genes in Breast Cancer

By Sai Saharsh Gumudavelly

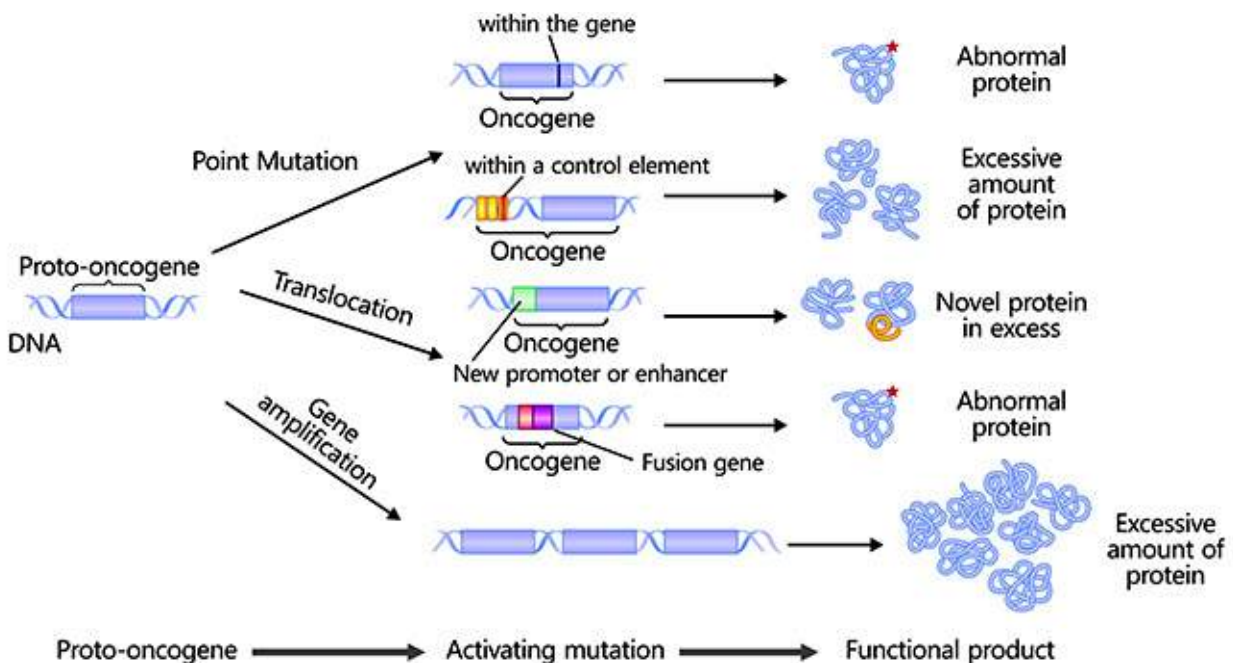
Cancer is a cellular malfunction where mutated cells grow uncontrollably and spread throughout the body. Cancer can originate in nearly any organ or tissue, making it one of the leading causes of death globally, impacting billions. In 2020, the World Health Organization reported approximately 10 million cancer-related deaths, underscoring its significance as a major global health issue. Risk factors for cancer are numerous, with lifestyle choices such as smoking and diet, exposure to environmental toxins, and genetic predisposition all playing key roles (World Health Organization, Cancer, 2022). Understanding these risk factors is crucial for both cancer prevention and treatment. However, despite increased awareness of its causes, cancer remains a prevalent and deadly disease, leaving a lasting impact on society. Mutations can also result in severe cardiovascular complications, such as myocardial infarctions and elevated blood pressure, and complications differ from mutation to mutation.

Cancer starts in the genes of cells and develops through a complex system of mutation. There are thousands of genes and gene-types, including the suppressor genes. Tumor suppressor genes encode certain proteins that regulate cell division. Tumor suppressor genes inhibit oncogenes, which drive abnormal cell proliferation, often leading to genetic mutations and cancer. These suppressor genes help control cell division, trigger apoptosis in mutated cells, and repair damaged DNA. By reducing oncogene protein production, suppressor genes slow down cell growth and the cell cycle. However, when tumor suppressor genes themselves mutate, they lose the ability to regulate cell division, allowing uncontrolled growth of mutated cells and tumor formation. Mutations in these genes can be caused by genetic or environmental factors. Smoking, a genetic history of cancer, inherited genetic mutations, exposure to high levels of UV radiation, and cancers from other sections of the body can all be possible causes of mutations in breast cancer cells (see Figure 1) (Yeager, 2021). These mutations include tumor suppressor genes.



(Yeager, 2021)Figure 1: This picture describes how TSG (tumor suppressor genes) act when they are mutated and the factors that can affect this.

Proto-oncogenes are another type of gene. These genes encode proteins to kickstart the production of new cells. They typically regulate cell growth and death but can lead to cancer when mutated. They are coded to introduce proteins that enable the reproduction and maintenance of cells. Proto-oncogenes kickstart the rapid production of proteins that inhibit the death of cells. In moderation and control, these cells can easily and safely control the creation and facilitation of tissues in organisms. However, proto-oncogenes can be easily mutated, and their protein pathway can be complicated. Proto-oncogenes can be mutated in several paths, and these mutations can either be benign or mutant. Mutations in DNA sequences, chromosomal rearrangements, or gene amplification can activate proto-oncogenes, transforming them into cancer-causing genes (see Figure 2) (CUSABIO team, *Oncogenes and cancer*, 2024). This can cause a chain reaction, which results in the eventual mutation of the cell.



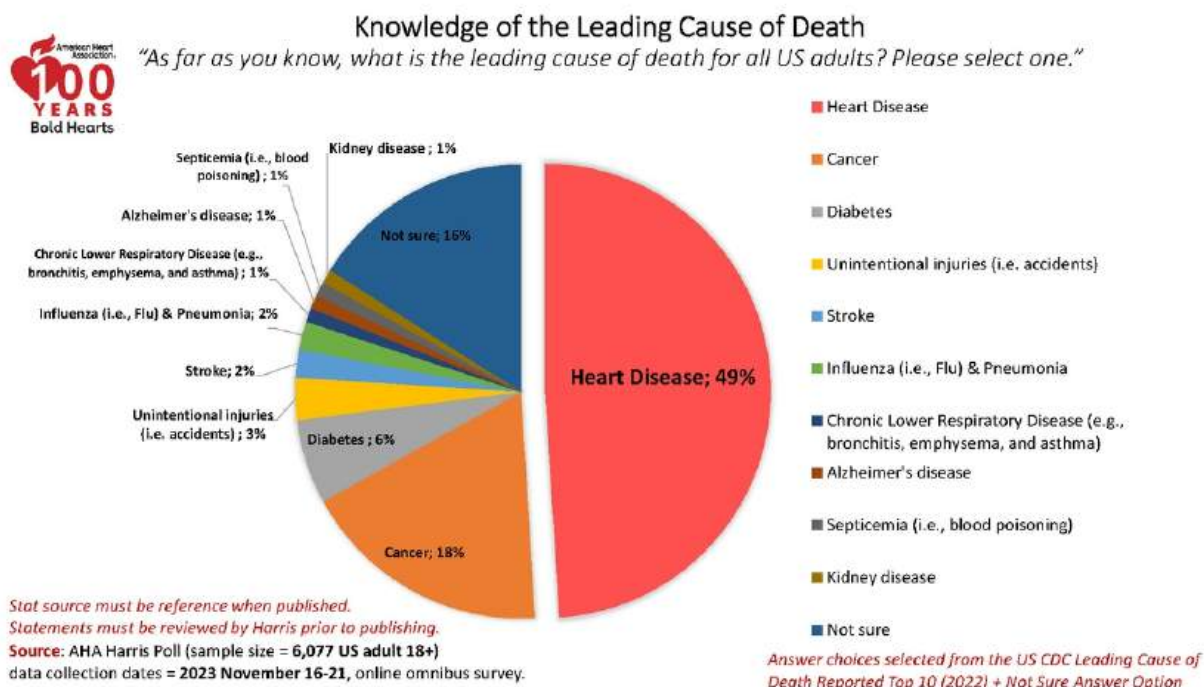
(CUSABIO team, *Oncogenes and cancer*) Figure 2: This diagram shows the numerous ways and their effects of unnecessary changes to DNA. The diagram shows the initiation of mutations and cellular/tissue malfunction.

This chain reaction can be devastating for tissues, organs, and even complete organisms. Mutations in proto-oncogenes classify them into oncogenes, which reproduce uncontrollably as their mechanism to reproduce is broken. This paper primarily discusses mutations that typically lead to the development of breast cancer and their added effects on the body. While breast cancer is one of the most researched and least fatal cancers, its effects on other areas of the human body can be detrimental and aren't as studied.

The genes that cause breast cancer - BRCA1/2, MYC, PI3K(CA), and CDH1 are 4 different genes that each encode proteins stimulating or subduing the cellular reproduction cycle, along with other functions. These genes are known to cause breast cancer in patients, but the

mutation of these genes can negatively affect numerous other bodily functions. BRCA1/2 are categorized as tumor suppressor genes that code against proto-oncogenes such as MYC and PI3K in the cell. MYC is an oncogene that motivates the production of cells and turns into a proto-oncogene when mutated. The PI3K/CA gene codes for the PI3K and AKT pathways, which also produce and regulate cellular division and function, while CHD1 primarily codes for the e-cadherin protein, which codes for tissue function and creation.

Mutations of the BRCA1/2, MYC, PI3K(CA), and CDH1 have severe effects on the body. One effect these mutations can incite is the development of cardiac issues and diseases. Cardiac diseases remain the leading cause of death worldwide (Dattani et.al, 2023) and are the cause of death for 49% Americans (see Figure 3) (U.S. Department of Health and Human Services, Causes and risk factors, 2023). While numerous factors play a role in creating and continuing cardiac diseases, such as environmental factors like lifestyle and diet (American Heart Association, More than half of U.S. adults don't know heart disease is leading cause of death, despite 100-year reign. American Heart Association., 2024), it is increasingly evident that genetic mutations also significantly impact cardiac function. While genetic mutation is not preventable, understanding the risks of these mutations can help future scientists develop new, focused avenues for treating cardiac diseases.



(U.S. Department of Health and Human Services, Causes and risk factors, 2023) Figure 3: The image discusses the percentage of causes of death in the United States. Heart diseases make up 49% of the deaths in the United States, taking up around ½ the chart.

Analyzing BRCA1 and BRCA2 and Their Effects on the Cardiovascular System

BRCA 1 and BRCA 2, otherwise known as breast cancer type 1 and breast cancer type 2, are categorized as tumor suppressor genes (ACOG, BRCA1 and BRCA2 mutations, 2022).

Tumor suppressor genes encode proteins responsible for regulating cell division and maintaining genomic stability. These genes counteract the activity of oncogenes, which promote abnormal and uncontrolled cell proliferation (Cooper, 2000). Excessive cell division often results in the production of genetically unstable cells, leading to mutations that contribute to the formation of tumors and cancer growth. Tumor suppressor genes inhibit oncogenes by reducing cell proliferation, allowing cells to undergo apoptosis or self-destruction, when mutations are detected. Tumor suppressor genes achieve this by limiting the production of oncogene proteins, decreasing cell cycle activity, and halting cell division. Additionally, tumor suppressor genes encode proteins that repair damaged DNA, further preventing the proliferation of mutated cells.

When tumor suppressor genes are mutated, the regulatory mechanisms that control the cell cycle are compromised, allowing unchecked progression. In the absence of functional tumor suppressor genes, there is no effective restraint on oncogene-driven cell division, resulting in the uncontrolled proliferation of mutated cells and the eventual formation of tumors. These mutations can be induced by various environmental and genetic factors, including smoking, a family history of breast cancer, inherited genetic mutations, prolonged exposure to ultraviolet radiation, or secondary cancers originating from other areas of the body (Feng et. al, 2018). For instance, BRCA1 and BRCA2 are well-known tumor suppressor genes that play a crucial role in counterbalancing the effects of proto-oncogenes like MYC and PI3K, both of which are implicated in breast cancer development.

To survive, a tumor exploits the host's resources, reproducing rapidly due to its genetic mutations. Since the loss of tumor suppressor function impairs the body's internal mechanisms to halt this proliferation, chemotherapy is employed to target rapidly dividing cells, including cancerous ones. In response to such oncogenic activity, therapeutic interventions are often administered to combat tumor growth and slow the progression of cancer. Chemotherapy, a non-specific treatment composed of alkylating agents, is commonly used for a variety of cancers, including breast cancer, and may be particularly effective for cases involving BRCA1/2 mutations. However, due to its non-selective nature, chemotherapy also affects other rapidly reproducing healthy cells, such as hair follicles and skin cells, leading to adverse effects in various parts of the body, notably the cardiovascular system (Hood, 1986).

Chemotherapy can also have deleterious effects on cardiomyocytes. As rapidly dividing cells, cardiomyocytes are susceptible to the mechanisms of chemotherapy, which is designed to target and neutralize fast-replicating cells. The cardiotoxicity induced by chemotherapy can be classified into four categories: direct damage, ischemia, arrhythmia, and pericarditis. The extent of cardiovascular damage depends on the specific chemotherapy regimen and its duration, with some cases resulting in devastating effects (Florescu et. al., 2013). For example, repeated administration of anthracyclines can cause irreversible cardiotoxicity, leading to myocardial fibrosis and the necrosis of cardiomyocytes. Anthracyclines produce free radicals—atoms with

an unpaired electron—that can bond aggressively with cardiac tissues. Free radicals disrupt essential regulatory processes such as calcium homeostasis and protein suppression. While certain treatments, such as monoclonal antibody (Mab) therapies, offer reversible effects, many chemotherapy-induced changes can be permanent (McGowan et. al., 2017). The cardiovascular issues' causes even permeate to just the tumor itself, adding even more risk to breast cancer patients.

Emerging research has revealed a link between BRCA1/2 mutations and increased susceptibility to cardiovascular diseases (CVDs) like ischemic heart disease (IHD), atherosclerosis, and other cardiovascular disorders. Clinical trials have shown that the overexpression of BRCA1/2 in patients with IHD points to an additional, distinct function of these genes beyond their role in cancer prevention (Saha et. al., 2009). These trials demonstrated a correlation between BRCA1/2 mutations and altered Nrf2 activity—a transcription factor that regulates cellular responses to oxidative stress. Nrf2 is vital in controlling the expression of detoxification and antioxidant-related genes and proteins, in addition to mediating anti-inflammatory processes. A reduction in Nrf2 activity due to BRCA1/2 mutations can impair the ability of cells to manage oxidative stress and inflammation, both of which are critical for cardiovascular health. Given that the cardiovascular system endures significant oxidative stress, the loss of this regulatory function can heighten the risk of developing various CVDs (Satta et. al., 2017).

Analyzing MYC and its Effects on the Cardiovascular System

MYC, or myelocytomatosis oncogene, is a collection of three genes (c-myc, l-myc, and n-myc). As noted in its name, MYC is a proto-oncogene, and proto-oncogenes play a huge role in cellular division and reproduction (GeneCards, MYC Gene - MYC Proto-Oncogene, BHLH Transcription Factor. Genecards suite, 2024). Proto-oncogenes are genes that encode proteins that are responsible for initiating cell production. Under normal conditions, proto-oncogenes regulate cell growth and apoptosis, maintaining a balanced process of cell reproduction and death (ACS, Oncogenes, tumor suppressor genes, and DNA repair genes. American Cancer Society, 2022). However, when mutated, proto-oncogenes can drive the development of cancer. These genes promote the synthesis of proteins that facilitate cell reproduction and prevent cell death. When functioning properly, they help regulate the formation and maintenance of tissues in organisms. Yet, proto-oncogenes are highly susceptible to mutations, and their protein signaling pathways are complex. Mutations can occur in several ways, including alterations in DNA sequences, chromosomal rearrangements, or gene amplifications, all of which have the potential to activate proto-oncogenes, transforming them into oncogenes—genes that drive the uncontrolled growth of cells. This activation can trigger a chain reaction, leading to the eventual mutation of cells. Mutations in proto-oncogenes disrupt their regulatory mechanisms, classifying them as oncogenes that drive uncontrolled cell proliferation (Cline, 1987). This uncontrolled growth can have devastating effects on tissues, organs, and entire organisms as mutated cells continue to multiply unchecked.

Mutation of the MYC gene not only disrupts the system of cellular reproduction but also has far-reaching effects on various physiological processes. The MYC gene plays a critical role in stimulating cell cycle progression by activating specific genes and encoding proteins that drive cell cycle advancement. When MYC undergoes mutation, it can contribute to the development of breast cancer, often necessitating aggressive treatments such as chemotherapy and anthracyclines (Dang, 2012). However, beyond their role in cancer, MYC mutations can have detrimental effects on the cardiovascular system, leading to a significant decline in a patient's overall health. A particular area of concern is vascular dysfunction, which can result from the overproduction, underproduction, or mismanagement of vascular endothelial cells. The vascular linings, composed of thin layers of specialized cells, regulate blood flow to various parts of the body. When these cells are mutated, they may either proliferate excessively, cease functioning, or malfunction, impairing blood circulation and preventing adequate oxygen and nutrient delivery to tissues. Studies in animal models, particularly in mice with a knockout or loss of MYC expression, have demonstrated weakened vascular functionality. The impaired and weakened vascular ability shown by mice in the study is correlated to mice with a malfunctioning MYC gene, or poor expression of the gene. This shows a negative correlation between a loss in MYC expression and animals (Liu et. al, 2024). This research establishes a negative correlation between the loss of MYC expression and vascular performance, further underscoring the gene's importance in overall cardiovascular health.

Another study has demonstrated a correlation between the loss of MYC expression, specifically c-MYC, and the accelerated aging, degradation, and increased dysfunction of endothelial linings and cells. Endothelial cells form the inner lining of blood vessels, and any impairment in their function can directly affect the integrity and performance of the entire vascular wall. This relationship underscores how the reduction of MYC expression can have a profoundly negative impact on vascular walls and linings. The same study also revealed that c-MYC plays a pivotal role in regulating the rate of senescence in vascular and endothelial cells (Florea et. al., 2013). Senescence, or cellular aging, influences the lifespan, functionality, and interactions of cells. Mutations in mechanisms that control cellular aging can severely compromise tissue health and functionality. Furthermore, the same study indicated that the loss of c-MYC expression disrupts normal cell growth and diminishes endothelial function. The phenotype exhibited by senescent cells can trigger a pro-inflammatory response, which, in severe cases, may lead to self-destructive processes within the body, such as complete loss of vascular function and internal bleeding. These findings highlight the critical connection between MYC function and vascular health, showing that MYC mutations could severely compromise vascular integrity.

Cardiac hypertrophy (HCM), the abnormal thickening of the heart's ventricular walls, presents another challenge associated with MYC expression. HCM can place significant strain on the heart, disrupting its ability to pump blood efficiently. Specifically, MYC has been linked to left ventricular hypertrophy, particularly in patients exhibiting other cardiac symptoms, such as chronic aortic regurgitation. Overexpression of MYC has been observed in patients with

moderate left ventricular thickening, which can be triggered by various factors, including MYC mutations. Interestingly, research has also found a lack of MYC expression in patients with severe hypertrophy, a phenomenon whose underlying mechanisms remain unclear. Growth factors, other genes, and protein pathways interacting with MYC may contribute to this variation in expression, which correlates with differing levels of myocardial hypertrophy (Napoli et. al, 2002).

Analyzing PI3K (specificity on PI3KCA) and its Effects on the Cardiovascular System

PI3K(CA) is classified as a proto-oncogene, responsible for encoding proteins that initiate and regulate cellular division. It plays a central role in activating the AKT pathway, which is crucial for the development and maintenance of tissues. This AKT pathway functions as a key regulatory mechanism in cellular processes, with the protein PI3K(CA) primarily responsible for its activation (Martini et. al, 2014). In addition to its regulatory role in cell division, PI3K(CA) also aids in glucose metabolism and inhibits apoptosis (cell death) under specific conditions, ensuring proper cell maintenance throughout the cycle. The gene tightly controls when a cell should reproduce and die, exerting strong regulatory influence over cellular production and maintenance (Hoxhaj & Manning, 2019). However, like all genes, PI3K(CA) is susceptible to mutation, which can lead to significant disruptions in its normal functions. One example of the impact of the PI3K(CA) mutation is Klippel-Trenaunay syndrome, a severe manifestation of PIK3CA-related disorders (Naganathan, 2023). Klippel-Trenaunay Syndrome, part of the broader group of PIK3CA-Related Overgrowth Spectrum (PROS) diseases, is characterized by abnormal and excessive growth of blood vessels and lymph nodes, leading to significant malformations (Klippel-Trenaunay syndrome: MedlinePlus Medical Encyclopedia, 2024). PROS encompasses a range of conditions caused by mutations in the PIK3CA gene, and these mutations often result in life-altering, severe complications. The malformations and overgrowth associated with these disorders can occur anywhere in the body, though they are typically localized, as the mutation does not affect every cell uniformly (Hammill & Nord, 2023). This distinction underscores the complexity and severity of diseases linked to PIK3CA mutations.

Klippel-Trenaunay is a specific subsection of these diseases. This syndrome confides in an overgrowth of vessels, lymph nodes, bones, and soft tissues. This syndrome is congenital, however. If present at birth, a patient would suffer from incorrect capillary formation, leading to internal bleeding, death of limbs, and extra tension placed on the cardiovascular system (see Figure 4) (Turner et. al, 2022) . Syndromes possible from this congenital syndrome include anemia, increased pain, blood clots, and even sepsis. These are all caused by the misdirection and incorrect flow of blood and result in overall negative health and degradation of the cardiovascular system. The syndrome is not genetic, however, and is a spontaneous mutation of the PI3K(CA) gene (Boston Children's Hospital, Klippel-Trenaunay syndrome. Klippel-trenaunay syndrome, 2024). The mutation of this gene causes a lack of control of the AKT and other pathways, leading to malformations seen in Klippel-Trenaunay's syndrome.



(Turner et. al, 2022)Figure 4: This image describes the negative effects seen by Klippel-Trenaunay syndrome. The incorrect formation of vessels leads to the loss of limbs, discoloration, and numerous vascular issues, as described in the picture. <https://artibiotics.com/blog/klippel-trenaunay-syndrome>

Analyzing CHD1/E-cadherin and its Effects on the Cardiovascular System

CDH1, a tumor suppressor gene, encodes for the protein E-cadherin, critical in maintaining cell-cell adhesion and ensuring proper cellular architecture within tissues. E-cadherin is primarily involved in forming junctions between cells. It acts as a gatekeeper to regulate cellular proliferation by maintaining contact inhibition—a process by which cells stop dividing when they reach a certain density (Bethesda (MD): National Library of Medicine (US), 2024). This function is crucial in preventing uncontrolled cell division, a hallmark of cancer development. Additionally, E-cadherin maintains tissue integrity by ensuring that cells remain attached, which prevents metastasis in cancerous cells. When CDH1 is mutated or its expression is lost, E-cadherin production is impaired, weakening cell adhesion and increasing the potential for tumor cells to invade surrounding tissues and spread to distant sites (Berx et. al, 1999). CDH1 and its product, E-cadherin, are vital for cell cycle regulation and maintaining tissue integrity, making their disruption a key factor in advancing many cancers, primarily breast cancer.

The mutation of CDH1 leads to the mutation and failure of E-cadherin, which can yield devastating results in the cardiovascular system. This failure disrupts the structural integrity of tissues and facilitates the invasion of cells, leading to cancer progression. Furthermore, this mutation can trigger angiogenesis, creating new blood vessels from pre-existing vessels. A study reveals how soluble E-cadherin, a protein found in high levels in the fluid surrounding ovarian cancer cells, strongly promotes angiogenesis, showing how a mutation of E-cadherin would

prove detrimental to the cardiovascular system. It does so through its release as a protein and exosomes—tiny vesicles that transport molecules between cells (Wong et. al, 2017) These exosomes bind to cadherin protein molecules, triggering a chain reaction that activates cellular pathways involved in cell growth and inflammation, linking cadherin to a significant cellular reproductive function and pathway. However, the impaired function of angiogenesis with an E-cadherin mutation can primarily increase the growth of blood vessels to cancerous mutations, allowing them to grow even faster and invade more organs (see Figure 5) (Stem Cell Regeneration Center, The Angiogenesis Process and Its Therapeutic Benefits, 2023). Impaired function could also lead to a complete shutdown, depriving new tissues and possibly leading to necrosis. A mutation in E-cadherin proteins can disrupt the junction connection between cells, deteriorating the communication and connection between tissue cells and majorly dismantling the creation of new vessels.

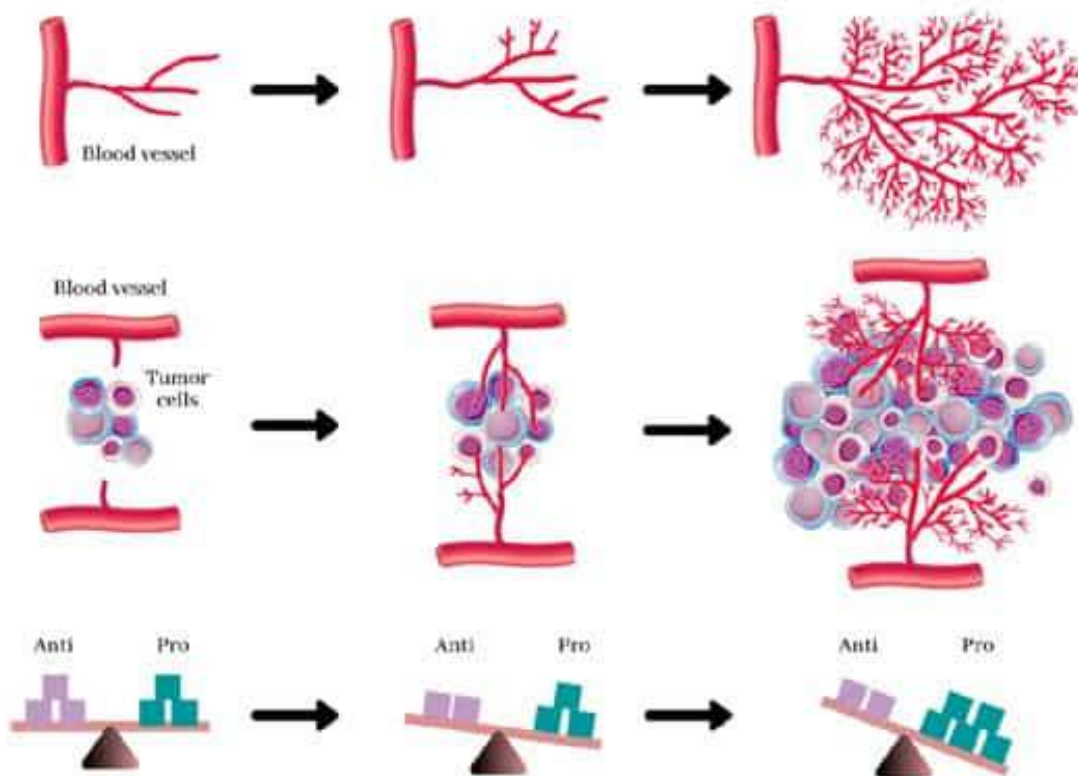


Figure 5: The image describes how blood vessels connecting to tumor cells increase their mass with the advanced progression of angiogenesis. The diagram also describes the branching process of angiogenesis, explaining how more blood reaches structures and tissue.

E-cadherin is also important for cell adhesion. Cell adhesion, or cell-to-cell attraction, is crucial for maintaining tissue integrity, particularly in heart muscle. Researchers led by Glenn Radice created transgenic mice that overexpress E-cadherin in the heart and found mice that have developed dilated cardiomyopathy. Mice that overexpressed e-cadherin also showed higher levels of DNA synthesis in heart cells, indicating the misexpression of e-cadherin and how it

triggers abnormal cell growth signals. Increased DNA synthesis indicates abnormal growth and synthesis of compounds, leading to incorrect and abnormal tissue formation (Ferreira-Cornwell et al., 2002). This suggests that inappropriate cadherin expression, a primary result of the mutation of CDH1, can cause cells, like cardiomyocytes, to behave abnormally and re-enter the cell cycle, disrupting normal tissue function.

Conclusion

Cancer starts with a simple mutation, but rapidly escalates the degradation of a host's body. Cancer is a multipronged attack on the body, aiming to siphon as many resources as possible from the host body. Cancer leads to the demise of multiple organs and organ systems, as shown throughout the paper. The relationship between gene mutations and their impacts on the cardiovascular system reveals a complex interplay between tumor suppressor genes and proto-oncogenes. Tumor suppressor genes typically regulate cell division by encoding proteins that restrict cell proliferation, initiate apoptosis, and repair damaged DNA, thus maintaining cellular health. When these suppressor genes mutate, cells are left vulnerable to unchecked growth, increasing cancer risk. On the other hand, proto-oncogenes promote the growth and reproduction of cells under normal conditions but can drive cancerous growth when mutated and activated as oncogenes. These mutations can disrupt the intricate cellular balance, profoundly affecting both cancer development and cardiovascular health, as illustrated by BRCA1/2, MYC, PBK(CA) mutations, and CDH1.

Mutations in BRCA1 and BRCA2, well-known tumor suppressor genes, are primarily associated with breast cancer. When mutated, these genes compromise the cell's ability to prevent abnormal growth, which can lead to tumor development and cancerous proliferation. Beyond this, BRCA1/2 mutations have been linked to increased cardiovascular risks, including ischemic heart disease and atherosclerosis. These genes influence Nrf2 activity, a transcription factor critical for combating oxidative stress and inflammation. A decline in Nrf2 activity due to BRCA mutations weakens cellular defenses against oxidative damage and inflammation, elevating the risk for cardiovascular diseases and exacerbating heart conditions under oxidative stress.

MYC, a proto-oncogene, encodes proteins essential for cell growth and division. When MYC mutates, it transforms into an oncogene, promoting the continuous division of abnormal cells. This unregulated proliferation negatively impacts the cardiovascular system, particularly vascular health. MYC mutations are associated with dysfunctional vascular endothelial cells, impairing blood flow and oxygen delivery. Additionally, MYC mutations can contribute to vascular aging, weakening blood vessel walls, and increasing susceptibility to inflammatory conditions. MYC has also been implicated in cardiac hypertrophy (HCM), where the thickening of heart walls disrupts efficient blood pumping, potentially resulting in chronic heart complications.

The PBK(CA) gene, another proto-oncogene, is integral to the AKT pathway, which

regulates cell division, glucose metabolism, and cell survival. Mutations in PI3K(CA) contribute to abnormal tissue growth, as seen in Klippel-Trenaunay syndrome, a condition characterized by extensive vascular malformations and overgrowth. Cardiovascular complications from PI3K(CA) mutations include internal bleeding, blood clots, and extreme vascular stress, conditions that place additional strain on the heart and circulatory system. The disorder of capillary and vascular formations resulting from these mutations can further lead to necrosis and sepsis, which can be fatal without medical intervention.

Finally, CDH1, a tumor suppressor gene, encodes E-cadherin, a protein vital for cell-cell adhesion and tissue integrity. Mutation of CDH1 disrupts E-cadherin production, weakening cell adhesion and enhancing the metastatic potential of tumor cells. Cardiovascular impacts of this mutation include impaired angiogenesis and compromised vascular integrity, increasing the risk for abnormal vascular growth and even internal bleeding in severe cases. Additionally, abnormal E-cadherin expression has been linked to dilated cardiomyopathy and dysfunctional cardiomyocytes, leading to heart enlargement and impaired pumping efficiency. These conditions strain the heart and may ultimately result in heart failure.

Works Cited

- World Health Organization. (2022, February 3). Cancer. World Health Organization.
- Yeager, C. (2021, March 4). 105 - tumor suppressor genes (w/ Pat Faughnan!). Curiosity: a Science Show.
- CUSABIO team. (n.d.). Oncogenes and cancer. CUSABIO.
- Dattani, S., Samborska, V., Ritchie, H., & Roser, M. (2023, December 28). Cardiovascular diseases. Our World in Data.
- U.S. Department of Health and Human Services. (2023, December). Causes and risk factors. National Heart Lung and Blood Institute.
- American Heart Association. (2024, January 24). More than half of U.S. adults don't know heart disease is leading cause of death, despite 100-year reign. American Heart Association.
- ACOG. (2022, December). BRCA1 and BRCA2 mutations. ACOG.
- Cooper, G. M. (2000). Tumor suppressor genes. *The Cell: A Molecular Approach*. 2nd edition.
- Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., Ji, X., Liu, W., Huang, B., Luo, W., Liu, B., Lei, Y., Du, S., Vuppalapati, A., Luu, H. H., Haydon, R. C., He, T.-C., & Ren, G. (2018, May 12). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and Molecular Pathogenesis. *Genes & diseases*.
- Hood, A. F. (1986, January). Cutaneous side effects of cancer chemotherapy. *Medical Clinics of North America*.
- Florescu, M., Cinteza, M., & Vinereanu, D. (2013, March). Chemotherapy-induced cardiotoxicity. National Library of Medicine.
- McGowan, J. V., Chung, R., Maulik, A., Piotrowska, I., Walker, J. M., & Yellon, D. M. (2017, February). Anthracycline chemotherapy and Cardiotoxicity. *Cardiovascular drugs and therapy*.
- Saha, T., Rih, J. K., & Rosen, E. M. (2009, April 11). BRCA1 down-regulates cellular levels of reactive oxygen species. *FEBS letters*.
- Satta, S., Mahmoud, A. M., Wilkinson, F. L., Yvonne Alexander, M., & White, S. J. (2017). The role of Nrf2 in cardiovascular function and disease. *Oxidative medicine and cellular longevity*.
- GeneCards. (2024, October 2). MYC Gene - MYC Proto-Oncogene, BHLH Transcription Factor. Genecards suite.
- ACS. (2022, August 31). Oncogenes, tumor suppressor genes, and DNA repair genes. American Cancer Society.
- Cline, M. (1987, September). The role of proto-oncogenes in human cancer: Implications for diagnosis and treatment. *International journal of radiation oncology, biology, physics*.
- Dang, C. V. (2012, March 30). Myc on the path to cancer. PubMed Central.
- Liu, W., Li, S., Yang, M., Ma, J., Liu, L., Fei, P., Xiang, Q., Huang, L., Zhao, P., Yang, Z., & Zhu, X. (2024, September 12). Dysfunction of calcyphosine-like gene impairs retinal

- angiogenesis through the Myc Axis and is associated with familial exudative vitreoretinopathy*. Genetics and Genomics.
- Florea, V., Bhagavatula, N., Simovic, G., Macedo, F. Y., Fock, R. A., & Rodrigues, C. O. (2013, September 6). *C-myc is essential to prevent endothelial pro-inflammatory senescent phenotype*. PubMed Central.
- Napoli, C., Lerman, L. O., de Nigris, F., & Sica, V. (2002, May). *C-Myc Oncoprotein: A dual pathogenic role in Neoplasia and cardiovascular diseases?*. Neoplasia (New York, N.Y.).
- Martini, M., De Santis, M. C., Braccini, L., Gulluni, F., & Hirsch, E. (2014, June 5). *Full article: PI3K/akt signaling pathway and cancer: An updated review*. Annals of Medicine.
- Hoxhaj, G., & Manning, B. D. (2019, November 4). *The PI3K-AKT network at the interface of oncogenic signalling and cancer metabolism*. PubMed Central.
- Naganathan, S., & Tadi, P. (2023, April 14). *Klippel-Trenaunay-Weber Syndrome*. StatPearls [Internet].
- A.D.A.M. Medical Encyclopedia [Internet]. Johns Creek (GA): Ebix, Inc., A.D.A.M.; c1997-2020. Klippel-Trenaunay syndrome; [reviewed 2023 Dec 31; cited 2024 Nov 04]; Available from: <https://medlineplus.gov/ency/article/000150.htm>
- Hammill, A. M., & NORD. (2022). *PIK3CA-related overgrowth spectrum - symptoms, causes, treatment: Nord*. Rare Diseases: PIK3CA-Related Overgrowth Spectrum.
- Boston Children's Hospital. (n.d.). *Klippel-Trenaunay syndrome*. Klippel-trenaunay syndrome.
- Turner, V. L., Kearns, C., Wattamwar, K., & McKenney, A. S. (2022, October 3). *Klippel-Trenaunay Syndrome | RadioGraphics*. RadioGraphics.
- MedlinePlus [Internet]. Bethesda (MD): National Library of Medicine (US); [updated 2020 Jun 24]. CDH1 gene; [updated 2017 Aug 1; cited 2024 Nov 4]; [about 6 p.]. Available from: <https://medlineplus.gov/genetics/gene/cdh1/>
- Berx, G., Becker, K.-F., Höfler, H., & van Roy, F. (1999, January 7). *Mutations of the Human E-Cadherin (CDH1) Gene*. Wiley Online Library: Mutation Update.
- Wong, S. H. M., Fang, C. M., Chuah, L.-H., Leong, C. O., & Ngai, S. C. (2017, November 20). *E-cadherin: Its dysregulation in carcinogenesis and clinical implications*. Critical Reviews in Oncology/Hematology.
- Stem Cell Regeneration Center. (2023, August 9). *The Angiogenesis Process and Its Therapeutic Benefits*. Stem Cell Regeneration Center News and Blog.
- Ferreira-Cornwell MC;Luo Y;Narula N;Lenox JM;Lieberman M;Radice GL; M. C., Luo, Y., Narula, N., Lenox, J. M., Lieberman, M., & Radice, G. L. (2002, April 15). *Remodeling the intercalated disc leads to cardiomyopathy in mice misexpressing cadherins in the heart*. PubMed.

Shifts in Migration Arrival Timing of Understudied South American Tern (*Sterna hirundinacea*) in the Falkland (Malvinas) Islands in Response to Climate Change

By Daphne Zhu

Abstract

As climate change continues to affect ecosystems globally, many migratory bird species are adapting to variability in climatic conditions by adjusting the timing of migration. However, this phenomenon remains underexplored in the Southern Hemisphere. In the Falkland Islands in the South Atlantic Ocean, the South American Tern typically arrives in austral spring and stays for the breeding season. In this study, the species' arrival timing in the Falkland Islands was analyzed with climatic conditions over the past ten years, revealing a shift toward earlier arrival in the year associated with increasing temperature in the Islands. The rate of this phenological shift in relation to the climatic ones should be closely followed into the future, since insufficient adaptation can lead to population decline. Further research on the South American Tern and other understudied species globally is critical to identifying species vulnerable to climate change impacts and informing conservation strategies.

1. Introduction

Anthropogenic climate change, driven by increasing greenhouse gas emissions, poses a major threat to species worldwide. One of the most profound impacts of climate change is its effect on the phenology—seasonal timing—of organisms globally (Inouye; Morten et al.). Migratory animals are particularly vulnerable due to their need to adapt to distant environments that may experience climate change at different rates. Their circannual rhythms are regulated by both internal mechanisms and environmental cues, such as photoperiod (i.e. length of day) or temperature (Hurlbert et al.).

Long-distance migrants often struggle to adjust their phenology to climate change compared to short-distance migrants, who are more attuned to local conditions. However, some studies indicate little difference in phenological shifts between the two groups (Da Silveira et al.; Hurlbert et al.; Zaifman et al.). Asynchronous migration timing can result in several challenges, such as unfavorable conditions upon early arrival, competition for territory or mates upon late arrival, and ecological mismatch with prey availability timing (Hurlbert et al.; Inouye). Species lacking sufficient phenotypic plasticity to adapt to a changing climate may experience increased risks of population decline and potential extinction (Zaifman et al.).

However, the phenological responses to climate change vary greatly due to differences in regional warming rates and each species' unique life histories and migratory behavior (Hurlbert et al.; Da Silveira et al.). Understanding how each species responds to climate change is critical for informing effective conservation and management strategies. Unfortunately, existing research is often biased, with a disproportionate focus on avian migration in the Northern Hemisphere (Da Silveira et al.). This leaves significant gaps in our understanding of how climate change affects

species in the Southern Hemisphere, particularly in marine ecosystems, which have received less attention compared to terrestrial ones.

Marine ecosystems face distinct climate change impacts, including decreased ocean productivity, shifts in species distributions and interactions, altered atmospheric patterns, and rising sea levels, which can reduce nesting habitat (Quillfeldt et al.). In addition, changes in sea surface temperature can alter prey availability during breeding seasons, directly impacting chick survival rates (Sydeman et al.). Climate change is one of the top three threats to seabirds, the most threatened of marine taxonomic groups, with 31% of species currently listed as Critically Endangered, Endangered, or Vulnerable (Dias et al.). As top predators, their populations serve as indicators of the overall health of the marine ecosystem (Sydeman et al.; Azuaje-Rodríguez et al.).

One such species is the South American Tern, *Sterna hirundinacea*, a poorly studied migratory seabird found along the Pacific and Atlantic coasts of South America. Despite its relative abundance and wide distribution across South America, the species faces many threats including egg harvesting, human disturbance, and Kelp Gull nest predation and is experiencing an overall decline in population (Faria et al.). Climate change may exacerbate this trend, particularly through sea level rise reducing nesting habitat and phenological mismatch that can reduce the species' overall fitness.

The South American Tern exhibits allochronic isolation, breeding in austral fall-winter in the north and spring-summer in the south (Azuaje-Rodríguez et al.). The breeding season corresponds to food availability, making it critical to align annual cycles with optimal conditions at breeding sites. Southern populations in particular perform long-distance migrations between breeding areas, in southern Chile and Argentina and the Falkland Islands, and overwintering areas, in regions like Uruguay and Southern Brazil (Azuaje-Rodríguez et al.). Specifically, the Falkland Islands are a key breeding site for South American Tern populations, with very little research for this species in this location.

This study focuses on the migration arrival of the South American Tern during breeding season in the Falkland (Malvinas) Islands, which occurs in austral spring-summer. By analyzing arrival timing using the online citizen science database eBird and correlating it with climatic indicators, this study aims to identify how climate change is influencing the phenology of this species. The study provides insight into its adaptive responses and potential conservation strategies.

2. Methods

This study focuses on the Falkland Islands (51° 48' S, 59° 31' W), an archipelago in the South Atlantic composed of two main islands and hundreds of small ones, covering approximately 12,200 sq. km. Data on the arrival dates of South American Terns (*Sterna hirundinacea*) were sourced from eBird, while meteorological data were sourced from the National Oceanic and Atmospheric Administration (NOAA). Linear regression models were

used to analyze the relationship between these environmental variables and the arrival date of the species in the Falkland Islands. All analysis was performed on R (version 4.4.1).

2.1 eBird Data

The online citizen science platform eBird was selected for its extensive and publicly accessible dataset, which supports robust analysis due to its comprehensive collection methods. Unlike other citizen science projects, eBird records provide detailed information about the observation process, increasing the reliability of the data. Data on the South American Tern (*Sterna hirundinacea*) observed in the Falkland Islands (51° 48' S, 59° 31' W) from 2014 to 2024 were downloaded from the eBird Basic Dataset (EBD) and Sampling Event Data (SED) from eBird (<https://ebird.org/data/download>).

To ensure data quality, the observations were filtered to include only those that had been verified and approved by regional reviewers, and only complete checklists were used in the analysis. To account for variation in observation effort, only checklists meeting specific criteria were included: 1) stationary or traveling, 2) less than 6 hours in duration and 10 km in distance, 3) travel speeds below 100 km/h, and 4) with no more than 10 observers (Mackey et al.). Detection/non-detection data were created by zero-filling the dataset, which enhances the robustness of the analysis compared to presence-only data.

The study focused on observations from 2014 to 2024, encompassing ten breeding seasons in the Falkland Islands in austral spring and summer. This time frame was chosen because eBird data has become more robust in recent years and is capable of reflecting responses to recent climatic trends (Zaifman et al.). Observations were grouped into 12-month segments from July of one year to June of the next to capture each breeding season comprehensively. No observations were recorded in the Falkland Islands between June 30 and July 5 of any year, as nearly all South American Terns breeding in the Falkland Islands migrate north to overwintering grounds.

2.2 Meteorological Variables

For this study, austral spring was defined as September 21st to December 20th (Brown et al.; Azuaje-Rodriguez et al., “Not going”). Five meteorological variables were selected for analysis: minimum spring temperature, mean spring temperature, mean spring precipitation, mean annual temperature, and mean annual precipitation. Each variable was considered for both the current year and the previous year to account for potential delayed responses to environmental changes (Møller et al.).

Meteorological data for the Falkland Islands were obtained from the National Oceanic and Atmospheric Administration (NOAA) database. Among the nine NOAA stations in the Falkland Islands, Mount Pleasant Station was chosen due to its complete and continuous dataset, including precipitation data, over the entire study period. Temperature data were verified with corresponding data from Stanley Station, the second most comprehensive station in the region.

2.3 Data Analysis

Firstly, to analyze temporal trends in the migration of South American Terns over the ten years (2014-2023), the date of the first observed tern in the Falkland Islands each year was selected to indicate migratory arrival. Linear regression was performed with the first observation date as the dependent variable and the respective year as the independent variable.

The number of days per season in which the species was observed varied over the study period, ranging from 4 to 36 days per season between 2014 and 2017, and from 78 to 166 days between 2018 and 2023. To account for this variation in data quality, a similar analysis was repeated for the period from 2018 to 2023. The results were consistent with the regression for the ten-year analysis, suggesting that the variation in data quality did not significantly influence the outcomes.

To address potential biases associated with using the first observation date (Garretson et al.), an alternative analysis was conducted using the median date of the first decile (10 %) of observations per season (abbreviated MD10). This date was calculated for observations made before Julian day 348 (December 14th), by which time the birds are expected to have completed their migration and begun breeding (Campioni et al. 4). This cutoff was determined by interpolating the median date of records before February, which was considered too late for new arrivals (Scolaro et al, Yorio et al. 300).

MD10 in 2014 was identified as an outlier based on the interquartile range from ten years' MD10 dates, and the year 2014 was excluded in all further analyses using MD10.

Spearman's rank correlation tests were used to assess the correlation between meteorological variables and year. Subsequently, linear regression was used to examine the relationship between each climatic variable and MD10. Both current and previous year meteorological data were included in these analyses to account for potential lagged effects of environmental conditions. Finally, as the last step in this analysis, a multiple regression model was developed to identify the most significant predictors of migratory arrival, with MD10 as the dependent variable and meteorological variables as the independent variables.

Results

Both the first observation date and MD10 date showed significant associations with the year. MD10 ranged from the latest of day 333 or November 29, occurring in 2014, to day 288 or October 15, occurring in 2021. Fig. 1 shows the decreasing trend in MD10 dates with a coefficient of -3.716, representing an earlier shift in migration arrival across the period, and a p-value of <0.05 indicating the statistical significance of this relationship between MD10 and the corresponding year.

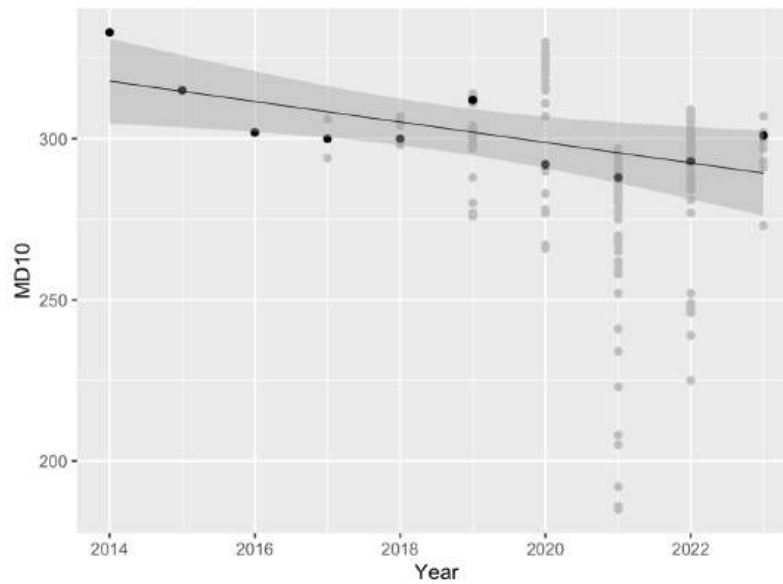


Fig. 1. The first 10% of South American Tern observations dates (2014-2023) in the Falkland Islands per season before the aforementioned cutoff date of December 14th, are represented in gray points. Black points and regression line correspond to the median date of these observations. The regression line has a constant of 6713.867, a coefficient of -3.176, and a p -value < 0.05 .

Among the meteorological variables, Spearman's rank test revealed a strong correlation between only mean spring temperature ($\rho = 0.63$, $p < 0.05$) and year (refer to Table 1). Though not statistically significant, the negative correlation between the two precipitation variables show a slight overall decrease in precipitation in the Falkland Islands in recent years, whereas the positive correlation in the two other temperature variables is consistent with the trend observed in mean spring temperature.

Table 1. Summary of Spearman's rank correlation tests between year and meteorological variables

	Minimum Spring Temperature	Mean Spring Temperature	Mean Spring Precipitation	Mean Annual Temperature	Mean Annual Precipitation
Year	0.536	0.627*	-0.467	0.500	-0.539

*: $p < 0.05$

Regardless of the correlation between the year and climatic variable based on Spearman's rank test, linear regression was performed between each climatic variable—using both current and previous year data—and MD10. Among these single-variable linear regression models, all three temperature variables using current-year data showed a significant association with MD10

($p < 0.05$). In contrast, none of the precipitation variables, nor any temperature or precipitation variables from the previous year, showed a significant association with MD10.

The three temperature variables were then combined to create a multiple linear regression model (refer to Table 2), with estimated coefficients of 1.163 for minimum spring, -6.318 for mean spring, and -12.732 for mean annual temperatures. In this model, a positive coefficient represents a later MD10 date with increasing temperature, while a negative coefficient represents a decrease (a shift earlier) in MD10 with increasing temperature. Although this model produced mixed positive and negative coefficients, only the variables with negative coefficients—mean spring and mean annual temperatures—were statistically significant. A similar process was used for a multiple linear regression model combining only these two variables (Table 3).

Table 2. Summary of Preliminary Multiple Linear Regression Model

This table includes only variables with significant results in the single-variable linear models, where MD10 is the dependent variable. A positive coefficient suggests that with an increase in temperature, there is a corresponding increase, or later shift, in the MD10 date; a negative coefficient represents a decrease or earlier shift.

Coefficients:	Estimate	Std. Error	p-value
(Intercept)	436.837	16.183	***
Minimum spring temperature	1.163	1.214	ns
Mean spring temperature	-6.318	1.286	**
Mean annual temperature	-12.732	1.632	***

Adjusted R-squared: 0.9682

***: $p < 0.001$

**: $p < 0.01$

ns: not significant

Table 3. Summary of Final Multiple Linear Regression Model¹

Coefficients:	Estimate	Std. Error	p-value
(Intercept)	423.6401	8.4168	***
Mean spring temperature	-5.3501	0.7887	***
Mean annual temperature	-12.1352	1.4978	***

Adjusted R-squared: 0.9687

***: $p < 0.001$

¹This table includes only variables with significant results from Table 2.

Finally, the equation of the last multiple linear regression model was plotted with the original data for mean spring temperature vs. MD10 (Figure 2), and mean annual temperature vs.

MD10 (not shown). The model provided a reasonable fit for the data in the mean spring temperature plot, while the mean annual temperature plot showed a noticeable skew.

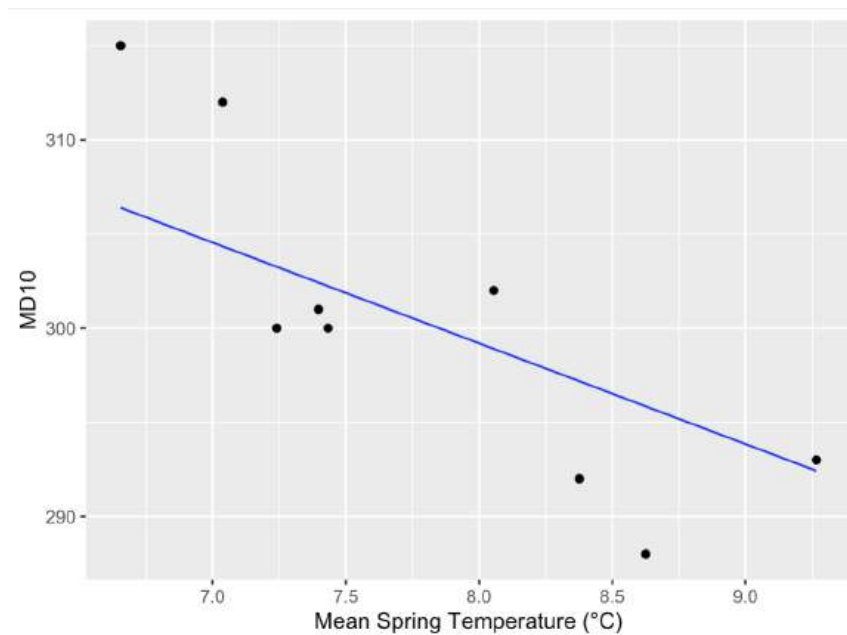


Figure 2. Data points for the mean spring temperature are plotted against MD10, while the blue line represents the equation of the multiple linear model presented in Table 2.

Discussion

The strong positive correlation between the mean spring temperature in the Falkland Islands and time (year) suggests that the impacts of climate change are already being reflected in rising temperatures in the region. Although Spearman's rank tests revealed insignificant correlations for the remaining temperature variables (i.e. minimum spring and mean annual), this may be due to greater variability in temperature during different seasons of the year, as well as in the single lowest temperature during spring compared to the spring average. The observed upward trends in minimum spring and mean annual temperatures still align with the trend in the mean spring temperature, further supporting the hypothesis of warming in the Falkland Islands. Not only does climate change affect temperature variables, but it could also impact weather events such as the El Niño Southern Oscillation or the Antarctic Oscillation (Quillfeldt et al.), which in turn could lead to greater variation in meteorological conditions in the Falkland Islands into the future.

The pattern visualized in Fig. 1 points to a gradual shift towards earlier spring migratory arrivals over the ten years. The association between higher temperatures and this shift may indicate the influence of climate change on the migratory phenology of the South American Tern. This is consistent with similar studies done on species such as migratory North American birds (Zaifman et al.), tropicbirds species (Campioni et al.), and Arctic terns (Møller et al.). Both the mean spring temperatures and mean annual temperatures showed a significant relationship but

differed in the estimated coefficients. This suggests that temperatures during other seasons, which are only reflected in the annual averages, may also have distinct impacts on migratory arrival timing and could be an important focus for future research. Interestingly, contrary to Møller et al.'s findings, this study found that the previous year's meteorological conditions did not have a strong effect on the arrival time of the terns. The birds may be more sensitive to current conditions, adapting to optimize short-term fitness.

Shifts in arrival date may affect the species' entire annual cycle, potentially causing misalignment with optimal environmental conditions throughout the year (Møller et al.). Future studies should explore additional phenological indicators including egg-laying dates, departure dates, and the frequency and duration of stopovers to gain a more comprehensive understanding of migration patterns (Zaifman et al.). Satellite tracking enables precise data collection on migration routes, while advanced climate models provide a broader understanding of migration timing within comprehensive climate patterns, beyond just temperature measurements.

There are, however, potential biases in the data that need to be considered. First, the eBird data used in this study is less standardized compared to data obtained from rigorous biological monitoring. Second, small sample sizes in earlier years could lead to results that are skewed and may not be comprehensive. Third, while the focus on meteorological data from the Falkland Islands was necessary, it may not fully reflect the shifts in arrival timing, as migration patterns are influenced by conditions in the birds' wintering grounds and stopover sites. Future studies should aim to include meteorological data from those regions to better understand the factors influencing migration phenology. Finally, although the data from Mount Pleasant station was cross-verified with another station, discrepancies among the nine stations of the Falkland Islands suggest that averaging data from multiple sources could reduce potential bias.

Despite these limitations, these findings contribute to our limited understanding of the species. The lack of research on this species, along with the relatively scarce eBird data in earlier years, highlights the urgent need for more focused studies. Academic research can highlight areas where climate change policies are needed, while continued community observations and regular monitoring enable local governments to track population status and implement optimal protection measures, ensuring the species' health into the future.

Conclusion

This study analyzed trends in the breeding season migration phenology of the South American Tern in the Falkland Islands with various climatic indicators. Findings revealed increasing temperatures in the region associated with earlier migration arrival timing across the last ten years. This points to the effect of climate change, which may be causing species to alter their phenological cycles in response to changing conditions. As the environment continues to change, understanding the phenological shifts in the South American Tern and other understudied species will be crucial for identifying those most at risk and informing future conservation efforts.

Works Cited

- Azuaje-Rodríguez, Roxiris A., et al. "Genetic and morphological variation of the poorly known South American Tern *Sterna hirundinacea*." *Marine Biology* 171.1 (2024): 21.
- Azuaje-Rodríguez, Roxiris A., Sofia Marques Silva, and Caio J. Carlos. "Not going with the flow: Ecological niche of a migratory seabird, the South American Tern *Sterna hirundinacea*." *Ecological Modelling* 463 (2022): 109804.
- Both, Christiaan, and Marcel E. Visser. "Adjustment to climate change is constrained by arrival date in a long-distance migrant bird." *Nature* 411.6835 (2001): 296-298.
- Bradley, Nina L., et al. "Phenological changes reflect climate change in Wisconsin." *Proceedings of the National Academy of Sciences* 96.17 (1999): 9701-9704.
- Brown, Jerram L., Shou-Hsien Li, and Nirmal Bhagabati. "Long-term trend toward earlier breeding in an American bird: A response to global warming?." *Proceedings of the National Academy of Sciences* 96.10 (1999): 5565-5569.
- Campioni, Letizia, Jeremy Madeiros, and Paolo Becciu. "Citizen science data reveal possible multi-decadal phenological changes in the arrival time of a migratory tropical seabird species at the breeding ground." *Marine Biology*, vol 170.10 (2023): 124.
- Campioni, Letizia, Jeremy Madeiros, and Paolo Becciu. "Early breeding site arrival of a migratory tropical seabird correlates with large-scale climatic phenomena in the North Atlantic." *Marine Ecology* (2023).
- Chen, Shuihua, et al. "Human harvest, climate change and their synergistic effects drove the Chinese Crested Tern to the brink of extinction." *Global Ecology and Conservation* 4 (2015): 137-145.
- Culp, Leah A., et al. "Full annual cycle climate change vulnerability assessment for migratory birds." *Ecosphere* 8.3 (2017): e01565.
- Da Silveira, Natália Stefanini, et al. "Future climate change will impact the size and location of breeding and wintering areas of migratory thrushes in South America." *The Condor* 123.2 (2021): duab006.
- Dias, Maria P., et al. "Threats to seabirds: a global assessment." *Biological Conservation* 237 (2019): 525-537.
- Faria, Patrícia J., et al. "Population structure in the South American tern *Sterna hirundinacea* in the South Atlantic: two populations with distinct breeding phenologies." *Journal of Avian Biology* 41.4 (2010): 378-387.
- Fracasso, Hélio AA, and Joaquim O. Branco. "Reproductive success of South American terns (*Sterna hirundinacea*) from Cardos Islands, Florianópolis, SC, Brazil." *Anais da Academia Brasileira de Ciências* 84 (2012): 527-536.
- Fracasso, Hélio Augusto Alves, Joaquim Olinto Branco, and Edison Barbieri. "A comparison of foraging between the South American and Cabot's Tern in southern Brazil." *Biota Neotropica* 11 (2011): 189-196.
- Garretson, Alexis, et al. "Citizen science data reveal regional heterogeneity in phenological response to climate in the large milkweed bug, *Oncopeltus fasciatus*." *Ecology and Evolution* 13.7 (2023): e10213.

- Grémillet, David, and Thierry Boulinier. "Spatial ecology and conservation of seabirds facing global climate change: a review." *Marine Ecology Progress Series* 391 (2009): 121-137.
- Hogan, Raissa Iris, Lia Jacobsen Prellvitz, and Carolus Maria Vooren. "Breeding biology of South American tern *Sterna hirundinacea* (Charadriiformes: Sternidae) on Deserta Island, Southern Brazil." *Revista Brasileira de Ornitologia* 18.3 (2010): 207-215.
- Hurlbert, Allen H., and Zhongfei Liang. "Spatiotemporal variation in avian migration phenology: citizen science reveals effects of climate change." *PloS one* 7.2 (2012): e31662.
- Inouye, David W. "Climate change and phenology." *Wiley Interdisciplinary Reviews: Climate Change* 13.3 (2022): e764.
- Johnston, Matthew Strimas-Mackey, et al. "Best Practices for Using EBird Data - 2 EBird Data." *Github.io*, 2023, ebird.github.io/ebird-best-practices/ebird.html.
- Møller, A. P., E. Flensted-Jensen, and W. Mardal. "Adjustment of the annual cycle to climatic change in a long-lived migratory bird species." *Current Zoology* 55.2 (2009): 92-101.
- Morten, Joanne M., et al. "Global warming and arctic terns: Estimating climate change impacts on the world's longest migration." *Global Change Biology* 29.19 (2023): 5596-5614.
- Portflitt-Toro, Matías, et al. "Breeding of the South American Tern (*Sterna hirundinacea*) on anchored boats in Coquimbo, northern Chile." *Ornitología Neotropical* 29 (2018): 187-191.
- Quillfeldt, Petra, and Juan F. Masello. "Impacts of climate variation and potential effects of climate change on South American seabirds—a review." *Marine Biology Research* 9.4 (2013): 337-357.
- Scolaro, José Alejandro, Sonia Laurenti, and Héctor Gallelli. "The Nesting and Breeding Biology of the South American Tern in Northern Patagonia (Nidificación y Biología de la Reproducción de *Sterna hirundinacea* en Patagonia Norte)." *Journal of Field Ornithology* (1996): 17-24.
- Seebacher, Frank, and Eric Post. "Climate change impacts on animal migration." *Climate Change Responses* 2.1 (2015): 1-2.
- Sydeman, William J., Sarah Ann Thompson, and Alexander Kitaysky. "Seabirds and climate change: roadmap for the future." *Marine Ecology Progress Series* 454 (2012): 107-117.
- Turella, Isadora Zanin, et al. "Breeding biology of swallow-tailed hummingbird (*Eupetomena macroura*) based on citizen science data." *Ornithology Research* 30.3 (2022): 181-189.
- Walker, Jacob, and Philip D. Taylor. "Using eBird data to model population change of migratory bird species." *Avian Conservation & Ecology* 12.1 (2017).
- Yorio, Pablo, et al. "Status and conservation of seabirds breeding in Argentina." *Bird Conservation International* 9.4 (1999): 299-314.
- Yorio, Pablo, et al. "Tourism and recreation at seabird breeding sites in Patagonia, Argentina: current concerns and future prospects." *Bird Conservation International* 11.4 (2001): 231-245.
- Zaifman, Jay, et al. "Shifts in bird migration timing in north American long-distance and short-distance migrants are associated with climate change." *International Journal of Zoology* 2017.1 (2017): 6025646.

Exploring Exoplanet Habitability: Equilibrium Temperature, Observational Constraints, and Implications for the Habitable Zone

By Ngoc Ha Bao Le¹ (Student) and Ana-Maria Piso-Grigore² (Mentor)

Abstract

In recent years, the study of exoplanets has emerged as a fascinating field of research. As an explanation, exoplanets hold the potential to unlock the pressing puzzle that has existed for more than 2000 years, such as “Do other Earths exist?” or “Do they have any sign of life?” [1]. It seemed that the question would never be answered and that the theory of exoplanets would only be a mere thing for science fiction and speculation for generations. However, with the advancement of technology, astronomers have effectively developed more techniques over the past 20 years, the majority of which depend on calculating the impact of orbiting planets on distant stars. We may uncover clues that bring us closer to answering one of humanity's most profound questions: “Are we alone in the universe?”. A key parameter in assessing the potential habitability of these distant worlds is their equilibrium temperature. Therefore, this paper aims to explain the exoplanet’s potential to support life by calculating the equilibrium temperature based on the formula I will derive and comparing those calculated results with confirmed data from NASA [2]. The analysis reveals the correlations between equilibrium temperature, the mass of host stars, and other factors that cause discrepancies beyond stellar irradiation, such as atmospheric composition and dynamics.

Introduction to exoplanet

Exoplanet formation

Exoplanets are planets outside our solar system, orbiting stars other than the Sun. They form through a process known as stellar formation, which involves the gravitational collapse of a large, rotating cloud of gas and dust. Molecular clouds are dense areas of spiral galaxies where exoplanet formation starts. These clouds contain various gaseous substances, including carbon dioxide, methane, water vapor, and organic molecules, but their main constituent is molecular hydrogen. Little solid dust grains, left over from earlier star generations, are dispersed throughout the gas. The collapse of an immense molecular cloud under its gravitational pull occurs due to a combination of factors, primarily gravity and the initial conditions of the cloud. The collapse of a vast molecular cloud under self-gravity is brought about by a combination of factors, including gravity and initial conditions [3]. The outcome is an acceleration in rotation similar to when a figure skater pulls his or her arms closer, resulting in them rotating faster. In this spinning cloud’s center, a protostar starts. This protostar is like the top’s central axis, which gains mass gradually as material falls inwards until it has gained enough mass to ignite nuclear fusion and become a star.

On the other hand, planets form via several processes like mutual gravitational attraction, van der Waals interactions, and electrostatic forces, which enable small ice and dust grains to

collide inside a gas and dust-made protoplanetary disk. The larger bodies (planetesimals) ranging in size from centimeters to kilometers assemble over time through collision between smaller grains or pebbles; moreover, planetesimals keep growing due to further collisions with each other or interaction among them by gravity until they end up creating protoplanets. Protoplanets are usually embryonic forms that have achieved substantial mass in most cases between a few kilometers and a few thousand kilometers across. The protoplanet increases its gravitational pull on the surrounding matter as it accumulates more mass due to gravitational attraction, which leads to more accretion. The runaway accretion phase occurs when gas and dust near the protoplanets fall towards them because they attract each other gravitationally and adhere to their surfaces [4]. At this point, a positive feedback loop results from its growing mass, amplifying its gravitational impact. This will be the case until the protoplanets reach a mass threshold where gravitational interactions become more disruptive than accretionary or until they run out of material in the protoplanetary disk. At this time, the protoplanets are now actual planets [5].

The significance of studying exoplanets

Studying exoplanets is vital for many science disciplines, but primarily for astrophysics, planetary science, and astrobiology. It reveals more about planetary formation and evolution. By comparing exoplanets to planets in our solar system, we gain insights into our origin and identify common patterns and significant differences, enabling humans to see certain universal principles that govern planetary systems. Such comparisons allow us to understand why our planet is unique, making it suitable for life compared to other bodies. Besides, understanding the process of planetary formation and evolution provides a theoretical testbed to examine how planets come into existence around infantile stars on disks called proto-planetaries. Observations from exoplanetary systems reveal various sizes and compositions as well as configurations in orbits, among others, that help us to understand how different processes related to coming up with planets can lead to various consequences because these are responsible for specifying the structure of planetary systems. The search for possible inhabitable worlds is one of the key motivations behind studying exoplanets. Scientists can then focus on these targets when searching for life beyond Earth by finding exoplanets within their star's habitable zone. Locating habitable exoplanets widens our scope regarding other locations where life might thrive apart from the Earth, and it also helps us look for biosignatures on such faraway spheres. [6,7]

Method

With the desire to search for extraterrestrial life, scientists aim to identify the environment that could potentially support life that we already have a mutual concept of or even discover new forms of life. Therefore, understanding the habitability of exoplanets is one of the major motivations for future observations and missions that could provide evidence of life beyond Earth.

The equilibrium temperature of exoplanets serves as an essential tool for habitability assessment. Generally, equilibrium temperature, a fundamental concept in various fields of

science, refers to the temperature at which a system reaches a balance between energy input and output. In thermodynamics, the equilibrium temperature of two objects in thermal contact is the temperature at which their heat exchange rates are equal. In astrophysics, the term “equilibrium temperature” applied to exoplanets means the theoretical temperature these planets would possess if they absorbed and emitted energy efficiently as a black body. It could be understood that it is the temperature at which the incoming solar energy from the planet’s host star balances with the energy radiated by the planet itself [8].

Imagine that a bright Sun is shining down on a patch of grass. The glowing Sun acts as the host star and emits sunlight, while the grass acts as a black body, absorbing the sunlight that heats it and emitting the light back into space. If the grass is too cold, it absorbs more sunlight, increasing its temperature. Conversely, if it's too hot, it emits more energy, decreasing its temperature. To maintain a balance, the grass's temperature must be at the same rate of absorbing and emitting sunlight.

$$T_p = T_{\odot} (1 - A)^{\frac{1}{4}} \sqrt{\frac{R_{\odot}}{2D}} \quad (1) \quad [9]$$

Where $T_{\odot}$ is the Kelvin temperature of the parent star, a is the adimensional Bond albedo, $R_{\odot}$ is in AU and not km or solar radii for the radius of the parent star, D is in AU for the semi-major axis, and T_p is the equilibrium temperature.

To begin with, one of a planet's complexities is its albedo. This is the fraction of sunlight reflected from the surface relative to the amount striking the surface:

$$A = \frac{\text{amount reflected}}{\text{amount incident}} \quad (2)$$

On Earth, the average albedo is approximately 0.3. This value accounts for various surface types (land, water, ice, and vegetation) and atmospheric effects. When estimating the equilibrium temperature of exoplanets, scientists often use Earth as a reference. Earth’s albedo contributes to this choice. Some planets, which are ice-covered worlds, may have higher albedos, while other rocky planets may have lower albedos. A perfectly reflective planet would have an Albedo of 1, while an utterly black planet would have zero albedo [10].

To find the equilibrium temperature, we equate the power supplied by the star or incoming solar radiation with the power emitted by the planet, the outgoing thermal radiation.

$$P_{\text{absorbed}} = P_{\text{emitted}} \quad (3)$$

As discussed above, planets are supposed to radiate as a black body, so we express the outgoing flux using the Stefan–Boltzmann law:

$$L_p = P_{\text{emitted}} = \sigma T_p^4 4[\text{area}] = \sigma T_p^4 4\pi R_p^2 \quad (4)$$

Where σ is the Stefan-Boltzmann constant with value of $5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$, $4\pi R_p^2$ is the area of the planet, L_p is the luminosity of the planet or the total energy output of the star per unit time (measured in Watt)

To calculate the incoming solar radiation, we need to calculate the star's luminosity first because the energy per unit time received by the planet is the star's luminosity. We consider a star as a black body when calculating the equilibrium temperature of an exoplanet. This simplification lets us focus on the energy balance between a planet and its parent star. Thus, the luminosity of the star is expressed:

$$L_{\odot} = \sigma T_{\odot}^4 4\pi R_{\odot}^2 \quad (5)$$

Assuming the planet absorbs all the radiation that falls on its surface, the total energy received per unit time by the planet is given by the equation:

$$P_{\text{absorbed}} = (1 - A) \frac{L}{4\pi d^2} \pi R_p^2 = (1 - A) 4\pi R_{\odot}^2 \sigma T_{\odot}^4 \left(\frac{R_p^2}{d^2}\right) \quad (6)$$

Since a fraction of incoming radiation is not reflected into space, $(1-A)$ represents the portion of energy absorbed by the planet's surface. If the planet reflects all incoming radiation, A would equal 0. If some of the energy is reflected ($A > 0$), then the planet's temperature will be $(1 - A)$ [9].

So now we equate the equations (5) and (6), we have:

$$\sigma T_p^4 4\pi R_p^2 = (1 - A) 4\pi R_{\odot}^2 \sigma T_{\odot}^4 \left(\frac{R_p^2}{d^2}\right) \quad (7)$$

Cancel the R_p and $4\pi\sigma$, we are left with:

$$T_p^4 = (1 - A) T_{\odot}^4 \left(\frac{R_{\odot}}{d}\right)^2 \quad (8)$$

Finally, take both sides to the fourth power of equation (8), and we have the exact equation (1).

Results

Applying the formula for exoplanets' equilibrium temperature, I analyzed the equilibrium temperatures of the 50 confirmed exoplanets and compared my results to the established data.

Measured by NASA equilibrium temperature (K) and Calculated equilibrium temperature (K)

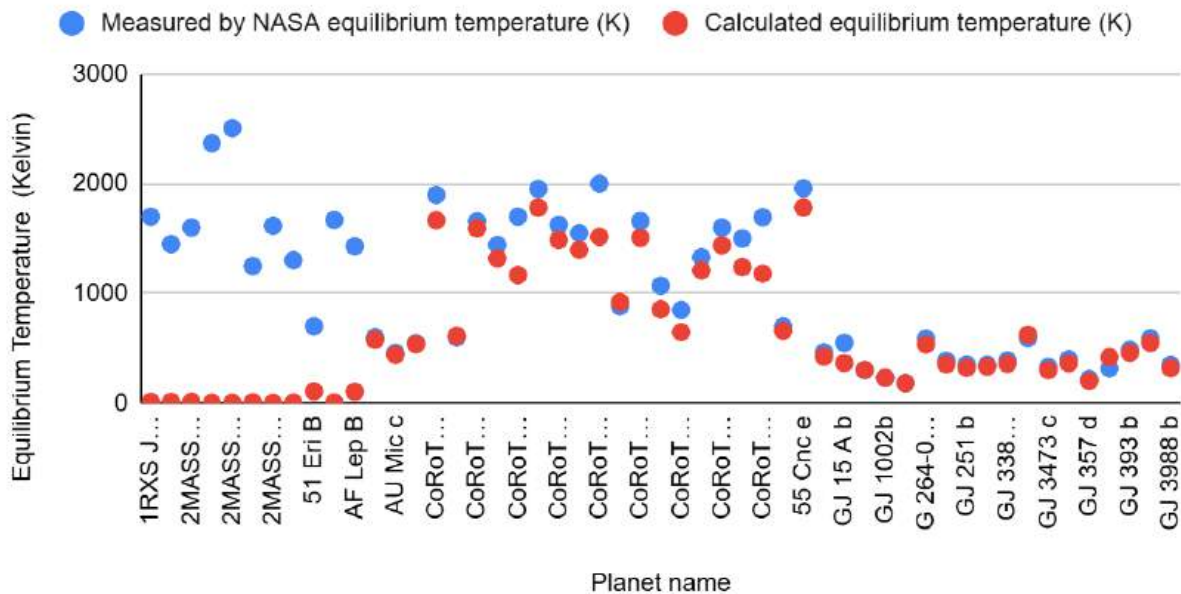


Figure 1: Comparison between equilibrium temperatures derived from theoretical calculations and the observed values obtained from the NASA exoplanet archive [2].

Looking at the chart, I observed some critical differences in the temperature discrepancies based on the detection method. Equilibrium temperatures calculated for exoplanets found through direct imaging are much lower than the observed temperatures in the data. Specifically, the first 11 exoplanets in the chart, measured by the detection method, have a calculated equilibrium temperature lower than 1000 kelvin, with the highest being 108.3 kelvin, much lower than their established figures, of which most are above 1000 kelvin. This difference is mainly because most directly imaged exoplanets are young gas giants that produce much of their heat, both from ongoing gravitational contraction and from the remaining heat of formation [11]. These internal heating sources are not taken into account by the formula for equilibrium temperature, resulting in significantly higher observed values than predicted.

However, slight discrepancies exist between calculated and observed temperatures for exoplanets discovered via the radial velocity method. Mainly, radial velocity detects older and more stable planets, often close to their stars, with predictable atmospheric conditions and equilibrium models [12]. Moreover, these exoplanets typically have less significant internal heating sources, unlike younger gas giants like those seen through imaging, whose calculated temperatures almost coincide with observations [12].

Similarly, transit photometry has minimal differences between the calculated and observed temperatures for such planets, besides its resemblance to radial velocity. Transit photometry usually finds planets orbiting closely around their parent stars, where atmospheric heat retention and redistribution are more significant [12, 13]. Nevertheless, this tends to be true because the dominant heat source is stellar irradiation, as opposed to any internal heat source affecting them.

Stellar Mass vs. Calculated equilibrium temperature (K)

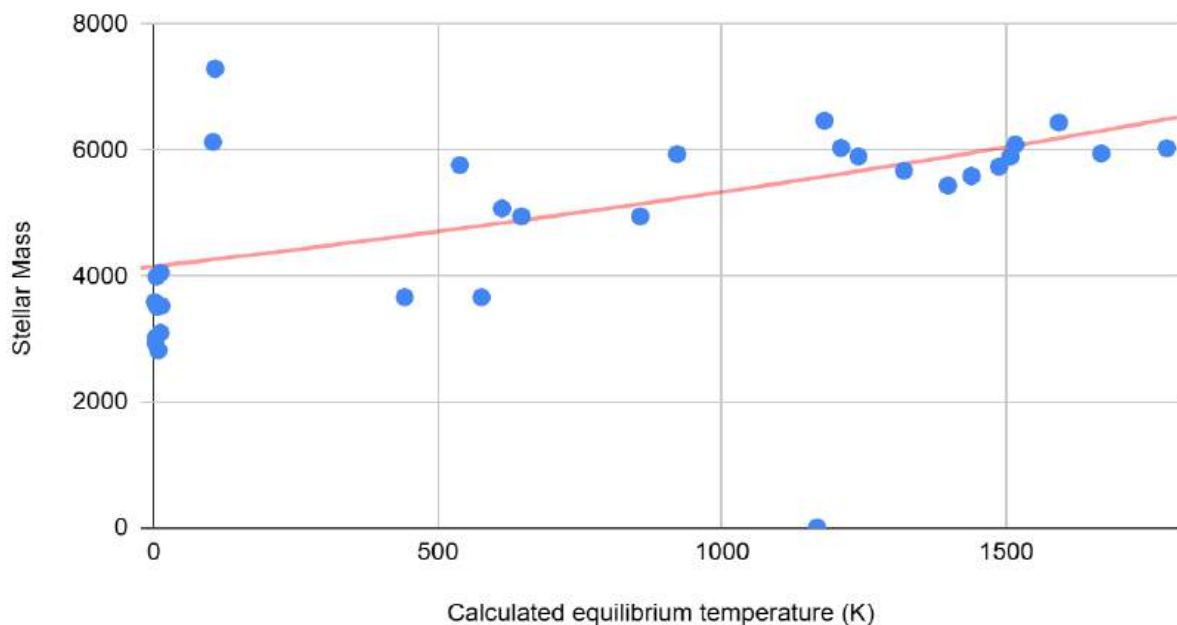


Figure 2: The correlation between the calculated equilibrium temperature and the mass of the star.

As we can see from Figure 2, there is a positive correlation between equilibrium temperature and stellar mass. As an explanation, the output energy of a star is related to its mass. According to the mass-luminosity relationship, a star's luminosity (or energy output) is proportional to its mass, raised to the power of approximately 3.5, so the more massive stars are, the more energy they emit. Another reason is that more giant stars will cause a stronger gravitational pull, so their attraction toward the stars is much greater. Planets orbiting more giant stars are closer to those larger stars, which emit more energy. However, this correlation is not entirely perfect because other factors, aside from the mass of stars, affect equilibrium temperature. For example, an exoplanet may have a star with a very high mass but may be in a very far orbit therefore, planet-star distance is considerable, and it will not receive much energy, whereas another exoplanet may have a very low mass host star, but might have a tight orbit therefore receives more energy from the star, so that means that the correlation between star mass and equilibrium temperature is not exactly perfect.

Discussion

However, as we can see, there is a discrepancy between the equilibrium temperature calculated based on the formula and the data reported by the NASA Exoplanet Archive [2]. The variation might have been explained by using atmospheric spectral retrieval techniques to extract database temperatures rather than those used in this document [14]. The database likely employs atmospheric models and spectral retrievals to estimate temperatures, while the simple formula for equilibrium temperature assumes no atmosphere.

In this method, the presence of different gases in an exoplanet's atmosphere affects how it absorbs and emits radiation. Some gases, like water vapor and carbon dioxide, absorb specific wavelengths of light. This absorption alters the energy balance, leading to temperature variations. Specifically, researchers have developed various retrieval suites for analyzing the exoplanet's spectrum [15]. Each suite uses its own forward models and parameter estimation techniques [16]. These models help build up simulated exoplanet spectra based on their atmospheric properties. They take input parameters, including the atmospheric composition, temperature profiles, and cloud properties, and predict the resulting spectrum (observed light) [16]. When an exoplanet transits in front of its host star, the star's light passes through the planet's atmosphere. The resulting spectrum shows how the atmosphere absorbs specific wavelengths of light. Scientists can infer the presence of different gases by comparing this spectrum to models.

Factors such as greenhouse effects and atmospheric composition change the equilibrium surface temperature of a planet, but these effects are not captured in the calculations. Exoplanets with dense atmospheres filled with greenhouse gases show consistently higher observed temperatures than their equilibrium temperatures. For instance, Venus has an atmosphere of thick carbon dioxide, resulting in an absolute surface temperature of around 737 K, much higher than its equilibrium temperature of about 227 K [17]. Exoplanets with substantial internal heat sources also have high observed temperatures. This is manifested in the case of gas giants such as Jupiter, whose unexpected internal heating due to gravitational contraction and other processes leads to elevated temperatures [11,18]. Atmospheric events such as strong winds, storms, and different compositions across various layers of the atmosphere can cause temperature variations that cannot be explained by simple equilibrium models like those used in meteorology [19]. Often, winds can blow air masses at different temperatures over long distances. For example, warm air from the tropics may move towards higher latitudes, while cold air from the poles may reach lower latitudes. This horizontal movement of air implies that models assuming static conditions do not capture temperature fluctuations. Winds vary the rate at which heat is removed or added to the ground. When wind speeds increase, there is more heat loss from surfaces, resulting in lower perceived temperatures. But during hot weather conditions, when it becomes windy, it mixes air that pushes warmer subtropical air down toward the surface. During cloud formation and precipitation, latent heat is released when water vapor changes into liquid droplets or solid ice crystals. Such release causes a warming of the surrounding environment through transient temperature anomalies that simplify equilibria do not account for. In addition, the jet streams - fast-flowing currents- affect weather systems at the surface, interacting with Rossby

waves and causing a thermal contrast [19]. These interfaces are also known as weather fronts and come about when two masses carry distinct properties.

Conclusion

In conclusion, this work has enhanced our knowledge of the habitable zone in exoplanets by demonstrating how well-established and tested physical principles can be utilized to predict equilibrium temperatures. As most exoplanets studied, this analysis indicated minimal variations between the calculated equilibrium temperatures and observed data. This implies that established theoretical frameworks are reliable tools for simulating atmospheric processes and climate on exoplanets. However, we found some instances where the observed temperatures differed significantly from the calculated ones. The fact that the calculations closely approximate observational data for most exoplanets strengthens the underlying theories, even though inconsistencies have been identified, which may necessitate further refining and expanding these models. Continued investigations in this direction will be crucial to discovering other planets' complex climates within or beyond the solar system.

Works Cited

- [1] Seager, Sara, Renée Dotson, and Houston. 2010. *Exoplanets*. Tucson: University Of Arizona Press; Houston.
- [2] "NASA Exoplanet Archive." 2019. Caltech.edu. 2019. <https://exoplanetarchive.ipac.caltech.edu/>.
- [3] Christensen, Majken Brahe Ellegaard. 2019. "How Do Stars Form?" *Frontiers for Young Minds* 7 (July). <https://doi.org/10.3389/frym.2019.00092>.
- [4] Lammer, Helmut, and Michel Blanc. "From disks to planets: the making of planets and their early atmospheres. An introduction." *Space Science Reviews* 214 (2018): 1-35.
- [5] "Planet Formation | Center for Astrophysics." n.d. www.cfa.harvard.edu.
<https://www.cfa.harvard.edu/research/topic/planet-formation>.
- [6] "Exoplanets | Center for Astrophysics." n.d. www.cfa.harvard.edu.
<https://www.cfa.harvard.edu/research/topic/exoplanets>.
- [7] "Your Guide to Exoplanets." n.d. The Planetary Society. The Planetary Society.
<https://www.planetary.org/worlds/exoplanets>.
- [8] Del Genio, Anthony D., Nancy Y. Kiang, Michael J. Way, David S. Amundsen, Linda E. Sohl, Yuka Fujii, Mark Chandler et al. "Albedos, equilibrium temperatures, and surface temperatures of habitable planets." *The Astrophysical Journal* 884, no. 1 (2019): 75.
- [9] 2021. Princeton.edu. 2021. <https://www.astro.princeton.edu/~strauss/FRS113/writeup3/>.
- [10] "Solar System Astronomy, Lecture Number 8." n.d. [Web.njit.edu](http://web.njit.edu).
<https://web.njit.edu/~gary/320/Lecture14.html>.
- [11] Thorngren, Daniel, Peter Gao, and Jonathan J. Fortney. "The intrinsic temperature and radiative-convective boundary depth in the atmospheres of hot Jupiters." *The Astrophysical Journal Letters* 884, no. 1 (2019): L6.
- [12] Mordasini, Christoph, G-D. Marleau, and P. Mollière. "Characterization of exoplanets from their formation-III. The statistics of planetary luminosities." *Astronomy & Astrophysics* 608 (2017): A72.
- [13] Las Cumbres Observatory. n.d. "Transit Method | Las Cumbres Observatory." [Lco.global](http://lco.global).
<https://lco.global/spacebook/exoplanets/transit-method/>.
- [14] Winn, Joshua N. "Transits and occultations." *arXiv preprint arXiv:1001.2010* (2010).
- [15] Barstow, Joanna K., Quentin Changeat, Ryan Garland, Michael R. Line, Marco Rocchetto, and Ingo P. Waldmann. "A comparison of exoplanet spectroscopic retrieval tools." *Monthly Notices of the Royal Astronomical Society* 493, no. 4 (2020): 4884-4909.
- [16] Welbanks, Luis, and Nikku Madhusudhan. "Aurora: a generalized retrieval framework for exoplanetary transmission spectra." *The Astrophysical Journal* 913, no. 2 (2021): 114.
- [17] Wenz, John. 2024. "Researchers Say This Is How to Find Alien Life Using the JWST." *Astronomy Magazine*. July 1, 2024.
<https://www.astronomy.com/science/greenhouse-gases-could-indicate-alien-life-if-we-ever-find-them/>.
- [18] "Thermal Radiation from Gas Giant Planets." n.d. Teach Astronomy - Thermal Radiation from Gas Giant Planets.

<https://www.teachastronomy.com/textbook/The-Giant-Planets-and-Their-Moons/Thermal-Radiation-from-Gas-Giant-Planets/>.

- [19] Showman, A. P., and T. Guillot. “Atmospheric Circulation and Tides of ‘51 Pegasus B-like’ Planets.” *Astronomy & Astrophysics*, vol. 385, no. 1, Apr. 2002, pp. 166–180, doi:10.1051/0004-6361:20020101.

A Review On the Effectiveness of Albumin Nanocarrier Delivery Vehicles When Combined With Paclitaxel in the Treatment of Metastatic Breast Cancer By Robin Chai

Abstract

Cancer treatment has been a subject of research that has been highly focused on in the past and present. Out of the many types of developing treatments, nanocarrier delivery systems are a particularly promising field as they provide a safer, enhanced and more targeted approach to the treatment of malignant tumours using chemotherapy due to their ability to target specific proteins present in cancerous cells. However, despite their extensive research, their use is still limited in oncology due to the specificity of each nanocarrier to a specific type of drug and a particular type of cancerous cells. This exploration seeks to investigate a promising form of nanocarrier delivery system used in conjunction with the drug paclitaxel in treating different types of cancer. This investigation will also explore whether new advancements could be made in nanocarrier delivery systems to result in a more effective and universal approach to improving the treatment of cancers.

1. Introduction

1.1 Cancer growth, Breast cancer and Triple-negative breast cancer

Cancer is often defined as the uncontrolled proliferation of cells caused by random mutations, which result from a variety of factors such as random mistakes, exposure to mutagens and genetics, leading to the suppression of genes which control apoptosis (cell death) or mutations of proto-oncogenes into oncogenes that cause increased cell replication (Cancer.gov, 2021). The rapid replication of abnormal cells leads to the formation of an abnormal mass called a tumour. Many tumours remain benign for their duration; however, a tumour becomes malignant when the cancerous cells begin the overconsumption of resources in the body initially meant for other healthy cells, weakening the immune system and the growth of different cells. Malignant tumours can develop blood vessels further to hijack the body's source of nutrients and resources if not identified early and eliminated. In the late stages of cancer, the tumour may develop a metastatic nature where cancerous cells can break off from the primary tumour and travel to other locations in the human body, developing into secondary tumours. Cancer remains one of the leading causes of death around the world, totalling over 9.7 million deaths in 2022.

Breast cancer is a specific type of cancer that begins its growth in the cells of the breast tissue. It is the second most common cancer after skin cancer diagnosed in the United States, most commonly presented in the form of lumps found in breast tissue and discolouration. Breast cancer is mainly caused by mutations in the BRCA1 and BRCA2 genes, and increased risks can be inherited from first-degree relatives. Breast cancer is the most common type of cancer that predominantly affects women (Petrucelli et al., 2022). In 2022 alone, breast cancer has recorded a total of 2.3 million cases and 670 thousand deaths. It often starts within the milk ducts and eventually spreads to surrounding breast tissue. Late-stage metastasis can occur, causing the

development of secondary tumours across the body. If left unchecked, it can prove to be fatal in many cases (Breast Cancer UK, 2024). Triple-negative breast cancer is a specific type of breast cancer that is also explored in this investigation. Triple-negative breast cancer is a type of cancer that is particularly aggressive and dangerous due to its unresponsiveness to oestrogen and progesterone and its inability to produce human epidermal growth factor receptor 2 (HER-2) proteins (all of which tell the cell and controls when they grow and divide). The lack of these growth factors means cancer cells do not need growth factors to divide and proliferate, thus increasing their ability to spread as well as their likelihood of recurrence after treatment (Dass et al., 2021).

1.2 Challenges with existing treatments of breast cancer

Existing cancer treatment involves a combination of chemotherapy, radiotherapy and surgery to eradicate cancerous tissues by killing cancerous cells to reduce the size of tumours or surgically removing the tumour. However, there are many drawbacks to each of the existing methods in the treatment of cancer that could be improved to increase the efficacy and safety of treating cancer patients. Radiotherapy and chemotherapy are effective at reducing tumour size – especially when cancer is metastatic, where surgical intervention is not possible – by destroying the genetic material of cancer cells and stopping their replication. However, radiotherapy and chemotherapy, although effective, also target healthy surrounding tissues, killing healthy cells and leading to a worsening immune system, as well as symptoms such as hair loss (alopecia), fatigue, weight loss and inflammation, all of which adversely affects the patient and hinders the treatment process (Majeed & Gupta, 2023). Surgical intervention is effective but not always plausible in cases of late cancer as tumours could be too large, dangerous or spread out to remove surgically. Surgical procedures are also an invasive form of treatment that may be difficult to recover from and can cause other risks, such as infection (Fido et al., 2023). The few methods above are often used in conjunction to achieve the ideal effect. The major problem that can be resolved in chemotherapy with the introduction of nanocarriers is the more targeted nature of the treatment, which can minimise damage to surrounding tissue, benefitting the patient's overall well-being. The area nanocarriers aim to improve is chemotherapy. A survey regarding the side effects of chemotherapy suggested that 43% of patients suffer from headaches, 90% suffer from fatigue, 95% from weakness, 76% experienced hair loss, 77% from nausea, 75% from vomiting, 40% from abdominal cramps, 14% memory impairment, 49% numbness, etc. (Altun and Sonkaya, 2018). This is due to the unspecific nature of chemotherapy, where it targets both healthy and cancerous cells, leading to extensive damage to healthy tissues (Cleveland Clinic, 2024). Nanocarriers aim to target cancer cells specifically, often through surface proteins that differ from normal cells. They are also capable of increasing drug uptake in patients and thus increasing the efficacy of chemotherapy in cancer treatment, making it worthwhile to explore how it can reduce side effects and improve the effectiveness of chemotherapy-based treatments.

1.3 Nanocarrier delivery systems as a promising solution

The most significant advantage of nanocarrier delivery systems in the treatment of cancer is their ability to target abnormal cancer cells and reduce damage to surrounding healthy tissue. As it has special receptors that are complementary to the receptors and proteins found on the surface of a specific cancerous cell. This means the drug will only be released when it comes in contact with the cancerous cell and will reduce the overall effect on neighbouring cells, thus mitigating the adverse effects usually associated with chemotherapy. This dramatically improves the quality of life for the cancer patient and the efficacy of the treatment. There are many types of nanocarriers, ranging from micelles to dendrimers (Edis et al., 2021). Each nanocarrier is created to be complementary with a specific type of cancer cell through its surface proteins or other receptors, which allows for the effective targeting of cancerous tissue. When used in conjunction with paclitaxel, an effective drug often used in the treatment of breast cancer, nanocarriers can significantly improve the success rate and mitigate the side effects of the treatment. One such promising nanocarrier delivery system is a protein-based one. In particular, the human serum albumin (HSA, Figure 1A) nanocarrier is the most commonly used with paclitaxel in cancer treatment and has been recently approved by the Food and Drug Administration (FDA). Thus, this review seeks to explore its effectiveness in enhancing breast cancer treatment through the analysis of multiple clinical trials.

1.4 Paclitaxel and the use of Human Serum albumin in delivery vehicle systems

Paclitaxel is a common chemotherapeutic drug used in the treatment of breast and ovarian cancers (Figure 1B). It targets the structural protein tubulin, composed of alpha and beta subunits forming the heterodimer. Tubulin can also form polymers called microtubules *via* lateral self-assembly between tubulin dimers to form protofilaments. Microtubules are responsible for many essential processes within the cell, such as mitosis, locomotion of cells, cell signalling, and intracellular transport of macromolecules. Paclitaxel functions by stabilising microtubules, thus reducing their dynamicity. This stabilisation means that microtubules do not function well and are not flexible during mitosis, reducing the rate of mitosis and leading to increased cell death. They also lead to increased production of microtubules, which are not functional, disturbing the normal dynamics of microtubules in the interphase of mitosis (Meštrović, 2010). Side effects of paclitaxel are consistent with most chemotherapeutic drugs. These include the risk of infection, bleeding and bruising, allergic reactions, muscle pain, numbness, hair loss, nausea and diarrhoea. It can also lead to urinary tract infections and hypertension. HSA nanocarrier delivery system is a promising solution to these side effects (Cancer Research UK, 2022).

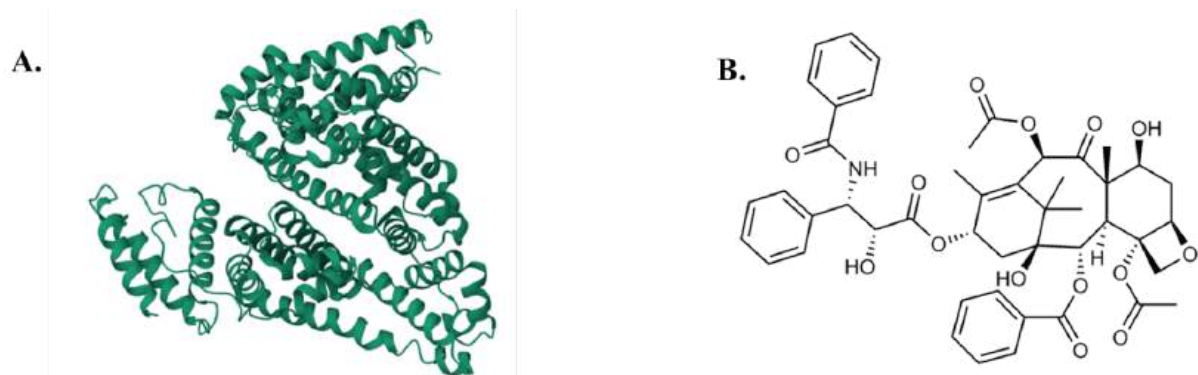


Figure 1. A. Chemical structure of paclitaxel. B. Structure of human serum albumin (HSA; PDB 1AO6).

HSA is a globular protein found in blood plasma composed of 585 amino acids. It also weighs around 66.5 kDa. A heart-shaped molecule is formed from three homologous domains. HSA is a versatile choice for synthesising nanoparticles as it is efficient at accumulation and tumour targeting when used in conjunction with cancer treatment drugs (Hassanin and Elzoghby, 2020). Enhanced tumour accumulation is due to the enhanced uptake passively mediated by the enhanced permeability and retention effect. Many tumours and cancerous cells overexpress the G-60 protein receptor and secrete the protein acidic SPARC, which are proteins that HSA can specifically bind to (Meng et al., 2022). This increases the uptake of nanoparticles and, by proxy, the carried drug and allows HSA to bypass the drug efflux process in cancer cells. HSA also shows promise in reducing cancer drug resistance in cancer cells beyond the basic delivery function (Figure 2A). Using HSA in composition with paclitaxel is shown to increase the efficacy of the drug's binding ability on the endothelium of tumours by almost 10 times and increase transport efficiency by approximately four times. Due to the abundance and overexpression of the two receptors mentioned above, G-60 protein and SPARC, on different types of cancerous cells and tumours, HSA shows promise in being a more universal nano-delivery system for multiple types of cancer (Figure 2B). Amino and carboxylic groups present in HSA allow for the isolation and targeting of ligands. Degradation and decomposition are also slow when using HSA as a delivery system due to its stable structure and nature. This allows for the delivery of more types of drugs and agents in cancer treatments. Protein nanocarriers such as HSA possess many advantages when compared to other forms of nanocarriers, including but not limited to reduced cytotoxicity and immunogenicity and higher biocompatibility and biodegradability. Their ability to both actively and passively target tumour cells through overexpressed receptors makes them more universal, versatile and effective. They are especially promising for their ability to overcome drug resistance in many types of cancers as well as their ability to control drug release (Hassanin and Elzoghby, 2020).

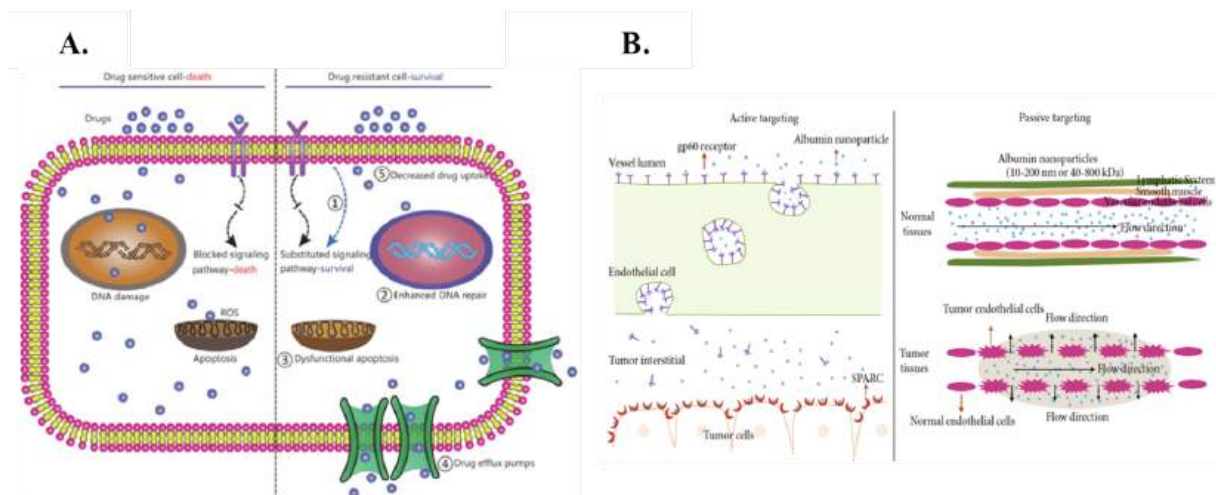


Figure 2. A. Representation of drug resistance and drug-sensitive cancer cells (source: Hassanin and Elzoghby, 2020). B. Mechanism of HSA nanoparticles in drug delivery and cancer treatment (source: Meng et al., 2022).

2. Methodology

This literature review paper focuses on the central analysis of whether the HAS nanoparticle delivery system effectively enhances paclitaxel and reduces the side effects in treating breast cancers. The analysis and conclusion will comprise data collected from other research papers and clinical trials conducted within the last 10 years. This will give a comprehensive overview of the modern applications of HAS and the general use of nanoparticles in the treatment of patients. Recently published papers and discoveries in the field of nanoparticles were used to evaluate and conclude whether the use of nanoparticles, especially protein-based nanoparticles, in the delivery system of chemotherapy and drugs is overall an effective and worthwhile improvement to modern and existing cancer treatment or whether investments could be made on other aspects of cancer treatment that could yield more benefits and results than nanocarrier delivery systems.

Due to the inability of running clinical trials to collect primary data on the efficacy of the HSA nanocarrier systems in cancer treatment, this paper relies on secondary and existing data. One limitation of using secondary data is that the quality and reliability of the data cannot be controlled. The data are extracted from credible, reliable research papers and clinical trials to circumvent this issue. These can be found in well-regarded sources such as peer-reviewed scientific journals and clinicaltrials.gov (see Works Cited).

3. Main Text

This section analyses clinical trials comparing the efficacy, safety and costs of the drugs Abraxane® and Taxol® (paclitaxel-based drugs). Abraxane®, also named nab-paclitaxel, is a chemotherapy drug that combines paclitaxel with HSA, acting as a nanocarrier. It is often used in the treatment of breast and ovarian cancers (EMA, 2009). Taxol® is the branded name of paclitaxel, and no nanocarrier delivery system has been implemented (Cancer.gov, 2006).

3.1 Comparison of safety and costs between Abraxane® and Taxol®

A study conducted in China in 2016 assessed the safety and efficacy of general paclitaxel compared to that of Abraxane® (paclitaxel delivered via HSA-based nanocarriers) to determine whether the introduction of specific delivery systems influences safety and effectiveness in the treatments of advanced metastatic breast cancers (Drantisaris et al., 2016). This study also analysed the cost-saving potential of introducing Abraxane® into the Chinese healthcare system. Thus, it is also essential to consider whether introducing Abraxane® can reduce costs and strain healthcare resources. In this study, a total of 12 patients were treated with Abraxane® whilst 15 patients were treated with Taxol®. These patients were selected to be comparable due to similar ages, metastatic sites, and performance status. Patients were tested for their respective drug responses, with acute reactions noted down. Resource costs include all the needed expenses for premedication, administration, labour costs, later checkups due to adverse reactions, etc. This study also accounted for the time required to administer the medication and treat adverse reactions.

Of the twelve patients who received Abraxane®, nine were administered doses at a weekly interval, whereas the remaining three were administered the drug every three weeks. For patients in the Taxol® groups, six were administered doses at weekly intervals and nine patients every three weeks. Through this, the safety of both drugs and the optimum interval can be investigated, which provides a more comprehensive overview of the safety of Abraxane® compared to Taxol®. It is important to recognise that the sample size of this study is limited and represents only a small proportion of Chinese patients. Therefore, its results may not be completely conclusive for all scenarios. However, it does provide important insight into how Abraxane® can improve cancer treatment and healthcare efficiency.

The study concluded that out of the patients who were given Taxol®, 5 out of 15 (33.3%) experienced acute adverse reactions to the drug, 3 of which required initiation of supportive interventions after visiting physicians. The percentage of acute adverse reactions associated with Abraxane® was much lower, with only one minor response out of 12 (8.8%) resolved with a temporary stoppage of administration. These results are confirmed with parametric statistical tests to confirm significant differences. Thus, it can be concluded that the use of HSA-based nanocarriers reduces the number of acute adverse reactions to paclitaxel and can better guarantee the safety of the patients being treated for late-stage breast cancer. Although the numerical difference did not represent a statistically significant difference in this study, it was without further investigation into the causes of acute responses (as this study focused more on the benefits of Abraxane® on healthcare resources). This value cannot be taken at face value and can only be hypothesised; further confirmation from other studies is awaiting.

The results also concluded a significant reduction in time and cost when Abraxane® was employed instead of traditional Taxol®. The time taken for administration, infusion and treatment was only 84 minutes for Abraxane®, whereas it was 282 minutes for Taxol® ($p < 0.001$). The low p-value suggests that the data collected is statistically significant and that the differences between the two groups are major/significant. However, the preparation time of Abraxane® was longer

than that of Taxol® (27 minutes vs 10.5 minutes). There was a considerable reduction in the infusion time needed for Abraxane® compared to Taxol® (83.4 minutes vs 281 minutes).

Furthermore, the study concluded that there was a significant reduction in costs to the healthcare service in terms of drug preparation and administration. In addition, there was a 50% reduced cost of premedication when using Abraxane® rather than Taxol®. This leads to the associated costs averaging around \$58.57 per dose for Abraxane® and \$254.14 for Taxol® ($p < 0.01$). The low p-value shows that the pairwise t-test shows the difference between the two groups is statistically significant. An average of \$586.71 could be saved per patient over a full cycle of weekly doses (three doses) and over \$3520 per patient over six treatment cycles if the taxable therapy was switched with the use of Abraxane®. This, thus, shows that using HSA-based nanocarrier delivery systems can reduce costs to resources and time due to increased efficacy/drug uptake and decreased side effects. This could potentially mean significantly reducing the pressure on healthcare costs, and resources can be redirected to more effective and efficient cancer care.

The study concludes that their findings show that Abraxane® is associated with a reduced adverse reaction rate compared to Taxol® (though not concrete) and a significant reduction in the costs of healthcare resources. The researchers claim that this can be beneficial to providing timely and effective care in the Chinese healthcare system due to the large number of breast cancer patients.

3.2 Investigation on the efficacy of paclitaxel bonded to HSA with normal paclitaxel in chemotherapy treatment of triple-negative breast cancer

A clinical trial aimed to investigate whether paclitaxel delivered by nanocarriers is more effective in chemotherapy treatment of triple-negative breast cancer than normal paclitaxel was published in 2020 (Yuan et al., 2020). Triple-negative breast cancer is a rare type of cancer associated with an accelerated rate of metastasis and tumour growth due to cancer stem-like cells (CSCs), which can replicate at a much faster rate compared to normal breast cancer cells as they lack receptors to oestrogen and progesterone and does not produce the HER-2 Protein. This means it does not need hormones to stimulate growth, allowing it to proliferate and spread at uncontrolled speeds. This makes it one of the fast-growing, dangerous and recurring cancers with limited forms of treatments. Paclitaxel has been identified as one of the most effective treatments for triple-negative breast cancer.

The study assessed both the number of cancer cells and the number of cancer stem-like cells when treated with either Abraxane® or Taxol® to identify whether the HSA nanoparticle delivery system seen in HSA is more effective at reducing the number of cancer cells but also the cancer stem cells, which can help reduce the chance of future resurgences or relapse after resolving the current tumour. This was conducted on lab rats to test the drug's efficacy. The study also collected data on cellular uptake of the paclitaxel in the two groups to assess the nanoparticles' effectiveness further.

The study identified that although Abraxane® was shown to have higher efficacy than Taxol® in treating metastatic breast cancer. The results showed that less Abraxane® was found in the blood plasma than Taxol® after the administration. This was initially attributed to a higher drug uptake into the tumour, but later investigations on triple-negative breast cancer cells (SUM149 cells) showed that accumulation of Abraxane® and Taxol® in the tumour cells are both similar in concentrations and viability of the drug in the tumour. Thus, drug accumulation and concentration in the tumour and blood cannot explain the higher efficacy of Abraxane® seen in clinical trials compared to Taxol®. However, this study concluded that both drugs effectively reduced tumour size and the number of cancer cells. Abraxane® managed to decrease the number of cancer stem cells, whereas Taxol® increased this number. This shows that Abraxane® is much more effective at long-term treatment and preventing relapses of triple-negative breast cancer. The efficacy against cancer recurrence was tested through transplantation of orthotopic breast cancer xenograft (1,000,000 SUM149 cells) mimicking triple-negative breast cancer cells in a mouse. Then, the mice were split into three groups: a control group treated with saline, a group treated with Taxol® and a group treated with Abraxane®. Data on their tumour size/volume and formation for both primary tumour sites and metastatic tumour sites were then tested over a few weeks to assess whether Abraxane® was more effective in the treatment of triple-negative breast cancer.

The data collected showed that the saline group experienced higher tumour growth in almost all categories compared to Taxol® and Abraxane®, confirming the effectiveness of paclitaxel with or without using HSA-based nanocarriers. Compared to Taxol®, Abraxane® did not show significant efficacy in controlling the growth of the primary tumour in these mice, where both Abraxane® and Taxol® showed similar trends of tumour growth. However, the same cannot be said for metastatic tumour sites where Taxol® showed a much higher tumour formation rate, whereas Abraxane® was able to significantly reduce the tumour formation rate and metastasis of the cancer. Most secondary cancer sites via metastasis in triple-negative breast cancer are caused by the cancer stem cells. The researchers of this investigation proposed that Abraxane® had a much better effect on reducing the number of cancer stem cells, whereas Taxol® focused only on shrinking already differentiated cancer cells.

To test this proposal, researchers conducted secondary implantation of dissociated single suspensions of the primary tumour into the mammary gland rates to evaluate further Abraxane® and Taxol®'s ability to eliminate cancer stem-like cells. The research discovered that Taxol® did indeed significantly increase the tumour regrowth rate in the secondary tumour and increased CSC frequencies by fourfold compared to the saline control group. Abraxane®, on the other hand, decreased the CSC frequency by twofold compared to the saline control group.

Abraxane® and Taxol®'s efficacy at eliminating aldehyde dehydrogenase cancer stem cells (ALDH+CSCs) was tested by detecting the aldehyde dehydrogenase 1 (ALDH1) present on ALDH+CSCs with staining of anti-ALDH1 antibodies. The percentage of CSCs in Taxol® treatment groups was 3 times higher than the saline control group, whereas there was an insignificant difference between the Abraxane® and control groups. The p-value of the pairwise

t-test is lower than 0.01, suggesting that the difference between Taxol® and the other groups is concrete and significant. Immunohistochemistry staining (IHC) also displayed results similar to those of the previous tests. The efficacy of the two drugs on another CSC found in triple-negative breast cancer- CSCs with cell surface adhesion receptor 44 and are missing cell surface adhesion receptor 24 (CD44+ /CD24-) are also tested *in vivo*, using a fluorescence-activated cell sorter (FACS). This showed that Abraxane® did not alter the number of this CSC compared to the control group, whereas Taxol® was shown to increase this type of CSC.

Further investigation was conducted on the efficacy of Abraxane® and Taxol®, compared in an adjuvant treatment where drug administration is stopped, and tumour regrowth is tested post-treatment. The frequency and number of CSCs will significantly impact the cancer regrowth rate, as a high CSC count will lead to rapid tumour regrowth, whereas a low CSC count will lead to slower regrowth. 50,000 SUM149 cells were transplanted into the mammary pads of three groups of rats. After four doses of treatment of both drugs, the rats were left to be observed for 120 days. The results displayed that both drugs showed significant efficacy in decreasing tumour size compared to the saline control group. As expected, Abraxane® displayed efficacy 3 times higher than Taxol® in reducing cancer growth.

To confirm whether the above results were significant, researchers conducted *in vitro* tests to verify that Abraxane® was more effective at reducing CSC formation than Taxol® could. This was performed on SUM149 cells via FACS. Although both Abraxane® and Taxol® showed no efficacy towards CD44+/CD24- CSCs (concordant with Abraxane® being similar to the control group in previous *in vivo* experiments), Abraxane® was shown to be 3-4 times more effective than Taxol® in the reduction of ALDH+ CSCs in an *in vitro* environment. Through this study, CSCs were identified to have self-renewal ability. Most cancer drugs, including paclitaxel, only inhibit primary mammosphere formation. However, if it does not penetrate and inhibit the secondary mammosphere formation, it is ineffective at decreasing the frequency of CSCs. Results confirmed that Abraxane® inhibited secondary mammosphere formation, whereas Taxol® increased secondary mammosphere formation.

Another aspect other than the increased efficacy of treatment was the increased uptake rate due to the introduction of HSA-based nanocarrier delivery systems in Abraxane®. Earlier investigations in this study did identify the fact that there was no significant difference between Abraxane® or Taxol® concentrations in tumour cells and even a lower Abraxane® concentration in blood plasma than Taxol® despite prior research showing that there is a higher uptake of Abraxane® compared to Taxol® in tumour cells. However, the ratio between Abraxane® concentration in the tumour and Abraxane® concentration in the blood plasma was significantly higher than that of Taxol®. This suggests a better uptake and absorption of Abraxane® compared to that of Taxol®.

The researchers thus tested *in vitro* the uptake rate of Abraxane® versus Taxol® in SUM149 cells. Results showed that Abraxane® was absorbed up to 8-15 times better than Taxol®. The enhanced penetration effect of Abraxane® in cancer-like stem cells is further reinforced by the data displaying enhanced uptake of Abraxane® compared to Taxol® in both

ALDH+ CSCs and ALDH- CSCs. P-values showed that the difference collected was significant enough to conclude that HSA-based nanocarriers are effective at enhancing the intracellular penetration and absorption of paclitaxel at 2 μ m concentration over the span of 60 minutes.

3.3 Limitations of the use of nanocarriers

Current clinical trials seem to suggest that nanocarrier delivery vehicles do have a positive effect on the treatment of breast cancer, with clinical trials showing that there were significant reductions in the numbers of deaths and some showed a decrease in tumour size, thus providing evidence to the efficacy of HSA-based nano delivery systems. However, current clinical trials need to provide insights into whether this type of nanocarrier can apply to other cancers. The most significant limitation associated with its nano-delivery systems is the specificity needed for every kind of cancer due to differing surface proteins. Although HSA, when used in conjunction with paclitaxel (Abraxane[®]), has shown efficacy, its breadth of use is still unclear. This does not mean that nanocarriers will be discounted for general cancer treatment in the future (Alshawwa et al., 2022).

4. Conclusions and Outlook

The use of HSA-based nanocarriers has shown to be effective at reducing side effects and increasing the efficacy of paclitaxel. This conclusion is derived from comparing clinical trials using paclitaxel alone and with HSA-based nanocarrier delivery systems. However, due to a lack of clinical trials directly comparing the two variables. The results may not be entirely consistent. It is also crucial to recognise that there are still limitations associated with using nanocarriers in cancer treatment, as mentioned above.

The development of intelligent nanocarriers that can adapt themselves to diverse types of cancerous cells based on their surface proteins and the surrounding environment can further contribute to the advancement of using nanocarriers as delivery vehicles. Recent studies suggest that when used in cancer treatments, smart nanoparticles respond to biological cues and internal and external factors such as pH, enzymes, optics, temperature, and magnetism. The ability to adapt according to the stimulus allowed them to be classified into intelligent systems. This field of innovation has progressed rapidly, especially with the introduction of Artificial intelligence. Although this prospect is not without its limitations and flaws, the potential of smart nanoparticles and their unique nature can help provide targeted, personalised and precise cancer treatment (Sun et al., 2023).

Works Cited

- Alshawwa, Samar Zuhair, Abeer Ahmed Kassem, Ragwa Mohamed Farid, Shaimaa Khamis Mostafa, and Gihan Salah Labib. "Nanocarrier Drug Delivery Systems: Characterisation, Limitations, Future Perspectives and Implementation of Artificial Intelligence." *Pharmaceutics* 14, no. 4 (April 18, 2022): 883. <https://doi.org/10.3390/pharmaceutics14040883>.
- Alqahtani, Fulwah Yahya, Fadilah Sfouq Aleanizy, Eram El Tahir, Hamad M. Alkahtani, and Bushra T. AlQuadeib. "Paclitaxel." In *Profiles of Drug Substances, Excipients and Related Methodology*, 44:205–38. Elsevier, 2019. <https://doi.org/10.1016/bs.podrm.2018.11.001>.
- Altun, İnsaf, and Alper Sonkaya. "The Most Common Side Effects Experienced by Patients Were Receiving First Cycle of Chemotherapy." *Iranian Journal of Public Health* 47, no. 8 (August 2018): 1218–19.
- Breast Cancer UK. (2024, October 4). Facts and figures | Breast Cancer UK. Breast Cancer UK. [https://www.breastcanceruk.org.uk/about-breast-cancer/facts-figures-and-qas/facts-and-figures/#:~:text=Global%20breast%20cancer%20statistics,deaths%20among%20women%20\(1\).](https://www.breastcanceruk.org.uk/about-breast-cancer/facts-figures-and-qas/facts-and-figures/#:~:text=Global%20breast%20cancer%20statistics,deaths%20among%20women%20(1).)
- Cancer.gov. (2021, October 11). What Is Cancer? Cancer.gov. <https://www.cancer.gov/about-cancer/understanding/what-is-cancer#:~:text=Cancer%20is%20a%20disease%20caused,are%20also%20called%20genetic%20changes>.
- Cleveland Clinic, C. C. M. (2024, October 21). Chemotherapy Side Effects. Cleveland Clinic. <https://my.clevelandclinic.org/health/articles/10257-chemotherapy-side-effects>
- Dass, Sylvia Annabel, Kim Liu Tan, Rehasri Selva Rajan, Noor Fatmawati Mokhtar, Elis Rosliza Mohd Adzmi, Wan Faiziah Wan Abdul Rahman, Tengku Ahmad Damitri Al-Astani Tengku Din, and Venugopal Balakrishnan. "Triple Negative Breast Cancer: A Review of Present and Future Diagnostic Modalities." *Medicina (Kaunas, Lithuania)* 57, no. 1 (January 12, 2021): 62. <https://doi.org/10.3390/medicina57010062>.
- Dranitsaris, George, Bo Yu, Leiping Wang, Weili Sun, Yan Zhou, Jennifer King, Satyin Kaura, Adams Zhang, and Peng Yuan. "Abraxane® versus Taxol® for Patients with Advanced Breast Cancer: A Prospective Time and Motion Analysis from a Chinese Health Care Perspective." *Journal of Oncology Pharmacy Practice* 22, no. 2 (April 2016): 205–11. <https://doi.org/10.1177/1078155214556008>.
- Edis, Zehra, Junli Wang, Muhammad Khurram Waqas, Muhammad Ijaz, and Munazza Ijaz. "Nanocarriers-Mediated Drug Delivery Systems for Anticancer Agents: An Overview and Perspectives." *International Journal of Nanomedicine* 16 (2021): 1313–30. <https://doi.org/10.2147/IJN.S289443>.
- Fico, Valeria, Gaia Altieri, Marta Di Grezia, Valentina Bianchi, Maria Michela Chiarello, Gilda Pepe, Giuseppe Tropeano, and Giuseppe Brisinda. "Surgical Complications of Oncological Treatments: A Narrative Review." *World Journal of Gastrointestinal Surgery* 15, no. 6 (June 27, 2023): 1056–67. <https://doi.org/10.4240/wjgs.v15.i6.1056>.

- Hassanin, Islam, and Ahmed Elzoghby. "Albumin-Based Nanoparticles: A Promising Strategy to Overcome Cancer Drug Resistance." *Cancer Drug Resistance*, 2020.
<https://doi.org/10.20517/cdr.2020.68>.
- Karami, Leila, Elham Tazikeh-Lemeski, and Ali Akbar Saboury. "Molecular Dynamics Simulation and Free Energy Studies on the Interaction of Salicylic Acid with Human Serum Albumin (HSA)." *Physical Chemistry Research* 5, no. 3 (September 2017).
<https://doi.org/10.22036/pcr.2017.63757.1315>.
- Majeed, H., & Gupta, V. (2023, August 14). Adverse Effects of Radiation Therapy. StatPearls - NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK563259/>
- Meng, Run, Huimin Zhu, Ziwei Wang, Shilei Hao, and Bochu Wang. "Preparation of Drug-Loaded Albumin Nanoparticles and Its Application in Cancer Therapy." Edited by Arivalagan Pugazhendhi. *Journal of Nanomaterials* 2022, no. 1 (January 2022): 3052175.
<https://doi.org/10.1155/2022/3052175>.
- News-Medical. (2023, May 13). How Paclitaxel works.
<https://www.news-medical.net/health/How-Paclitaxel-Works.aspx#:~:text=Paclitaxel%20stabilizes%20microtubules%20and%20reduces,tubulin%20and%20overproduction%20of%20microtubules.>
- Petrucelli, N., Daly, M. B., & Pal, T. (2023, September 21). BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. GeneReviews® - NCBI Bookshelf.
<https://www.ncbi.nlm.nih.gov/books/NBK1247/>
- Triple-negative Breast Cancer | Details, Diagnosis, and Signs. (n.d.). American Cancer Society.
<https://www.cancer.org/cancer/types/breast-cancer/about/types-of-breast-cancer/triple-negative.html>
- Sun, Leming, Hongmei Liu, Yanqi Ye, Yang Lei, Rehmat Islam, Sumin Tan, Rongsheng Tong, Yang-Bao Miao, and Lulu Cai. "Smart Nanoparticles for Cancer Therapy." *Signal Transduction and Targeted Therapy* 8, no. 1 (November 3, 2023): 418.
<https://doi.org/10.1038/s41392-023-01642-x>.
- Yuan, Hebao, Hongwei Guo, Xin Luan, Miao He, Feng Li, Joseph Burnett, Nathan Truchan, and Duxin Sun. "Albumin Nanoparticle of Paclitaxel (Abraxane) Decreases While Taxol Increases Breast Cancer Stem Cells in Treatment of Triple Negative Breast Cancer." *Molecular Pharmaceutics* 17, no. 7 (July 6, 2020): 2275–86.
<https://doi.org/10.1021/acs.molpharmaceut.9b01221>.

From Lev to Euro: Prospective Benefits of Eurozone Integration for Bulgaria

By Maria Kamenova

Introduction

The Report on Economic and Monetary Union in the European Community, also known as the Delors report, was submitted in April 1989 by the committee led by Jacques Delors, hence the informal name of the report. With the Delors report, the Economic and Monetary Union (EMU) foundations were laid, and the goal of creating a common currency (the euro) was officially set. The EMU consists of three stages. From 1990 to 1993, the first stage focused on “a complete removal of physical, technical and fiscal barriers within the Community” (European Commission, 1989). Its goal was also to “strengthen economic and fiscal policy coordination” among EU member states (European Commission, 1989). The second stage, from 1994 to 1998, established the European System of Central Banks (ESCB) and “the transition from the coordination of independent national monetary policies by the Committee of Central Bank Governors in stage one to the formulation and implementation of a common monetary policy by the ESCB itself scheduled to take place in the final stage” begun (European Commission, 1989). When a country enters the third stage of the EMU, starting from 1999, it officially adopts the euro: “In the course of the final stage, the national currencies would eventually be replaced by a single Community currency” (European Commission, 1989). In this sense, entering the third stage of the EMU is equal to becoming part of the Euro Area, also known as the eurozone (EZ).

A country must fulfill all euro convergence criteria defined within the Treaty of Maastricht from 1992 to enter the third stage. The euro convergence criteria consist of four main requirements: price stability—“this will be apparent from a rate of inflation which is close to that of, at most, the three best performing Member States in terms of price stability,” (European Union, 1992), in this sense, the inflation rate of a given Member State must not exceed 1.5 percentage points above the rate of the three best-performing Member States; sound and sustainable public finances—“this will be apparent from having achieved a government budgetary position without a deficit that is excessive,” (European Union, 1992), to be more specific, government deficit must not exceed 3% of GDP and public debt must not exceed 60% of GDP or be on a declining path; exchange-rate stability—“the observance of the normal fluctuation margins provided for by the Exchange Rate Mechanism of the European Monetary System, for at least two years, without devaluing against the currency of any other Member State,” (European Union, 1992), and long-term interest rates—“the durability of convergence achieved by the Member State and of its participation in the Exchange Rate Mechanism of the European Monetary System being reflected in the long-term interest rate levels,” (European Union, 1992), specifically, the long-term interest rates must not be higher than 2 percentage points above the three best-performing Member States.

Each country that joined the EU after 1999 must adopt the euro and fulfill all four criteria to join the eurozone. As of 2024, Bulgaria and Denmark are the only countries that are part of ERM II. Denmark, which joined the EU in 1973, negotiated an opt-out in 1992 after a

referendum on the Maastricht Treaty and is thus not obligated to introduce the euro. A referendum on reversing this opt-out was held on 28 September 2000, resulting “in a narrow, but still comfortable majority for continued non-participation in the euro-zone. As in earlier referenda concerning the EU, the turnout was high, at 88.1%, and 53.2% voted “no” and 46.8% “yes” to joining the euro,” (Carsten Jørgensen, 2000). On the other hand, Bulgaria joined the EU on January 1, 2007, and is required to adopt the common currency.

The European Commission’s Convergence Report 2024 suggests that Bulgaria is on the right path to becoming a member of the eurozone but is still not ready to introduce the euro: “Bulgaria is the only country that fulfils all but one criterion and where national legislation can be considered to be compatible with the rules of the Economic and Monetary Union,” (The Directorate-General for Economic and Financial Affairs (DG ECFIN), 2024). The only criterion that Bulgaria does not meet is price stability. In May 2024, Bulgaria's 12-month average inflation rate was 5.1% — “well above the reference value of 3.3% for the criterion on price stability” (the ECB’s Directorate General Economics, 2024).

It is also important to note the impact of the euro currency on different Member States, especially those in Eastern Europe which have joined in the past decade. In the case of Latvia, which adopted the euro in 2014, it is found that the country has benefited in many ways, such as “enhanced competition, bank intermediation, and the low-interest rate environment are bringing immediate and long-term benefits to the economy” (Gatis Eglitis, Radoslav Krastev, Christian Weise, 2014). Similarly, in Lithuania, which became a Member State in 2015, “after the euro adoption, the average wage and pension growth accelerated in Lithuania, far outpacing price inflation” (Linas Jurkšas, 2021). In Estonia, a eurozone member since 2011, the experienced “economic growth was 3.2 percentages in 2012” (Thomas Schneider, 2013).

Nevertheless, heated discussions continue in Bulgaria among politicians, economists, and the public regarding whether the country should introduce the euro as a common currency. With radical ideas such as “Vazrazhdane's call for a referendum that could rule the postponement of the euro adoption at least until 2043” (Kapital, 2023), there is still no clear answer to the many questions that are raised every day among the Bulgarian society regarding the eurozone integration. Moreover, there is not much information on the topic, as there is still no set date for adopting the euro. Society is divided over euro adoption, with 54 percent in favor and 44 percent against as of 2021, according to the study “Introduction of the euro in the Member States not yet having adopted the common currency” conducted by Eurobarometer (2021).

This study aims to answer the question of whether Bulgaria should join the eurozone by presenting evidence about the adoption of the euro in Bulgaria and adding to the existing body of literature on the topic, particularly because there is still limited evidence for Bulgarian society. Current evidence on eurozone integration, even though extremely limited, suggests that “the expected overall effect is strongly positive, especially for the case of Bulgaria” (Georgy Ganev, 2009) and “its implementation may not be ‘at any cost’” (Gechev, G., Beev, I., Hristozov, Y., 2020). The point of this paper is to provide clarity to both the public and policymakers and raise awareness about the potential eurozone integration. This is executed by conducting longitudinal

analysis, comparing non-eurozone Member States to countries that adopted the euro, and the main variables (or, in this sense, benefits) are the lower interest rates, higher levels of foreign direct investments (FDI), higher numbers of tourists in the countries that are part of the eurozone. The elimination of money exchange is also presented using a different approach. An important advantage also discussed is the possibility of emergency support by the ECB and better overall regulation hence lower corruption rates.

Methodology

Data & measures

For this comparison, I used three main interest variables: interest rates, foreign direct investments (FDI), and the number of tourists in each respective country. All data was obtained from the World Bank.

Comparing FDI among all EU Member States is crucial for recognizing the benefits of the eurozone integration: “Information on its [FDI’s] evolution over time provides an indication of an EU Member State’s external competitiveness and the effectiveness of economic, trade and investment policies” (European Commission, 2021). Measuring the foreign direct investments and their changes over time is a beneficial way of detecting the changes in a particular country’s economy: “these changes may be linked to global economic fortunes, sector-specific developments, or the individual situations of enterprises considering foreign investments” (European Commission, 2021).

As it is well known, interest rates “influence borrowing costs and spending decisions of households and businesses” (Federal Reserve Board, 2024). Low interest rates stimulate people to invest in mortgages and borrow money, encouraging big purchases and investments. It can also signify a stable, attractive economic environment as central banks often maintain lower interest rates to encourage borrowing and investments.

It is important also to mention the relationship between interest rates and bonds. Bonds have an inverse relationship to interest rates meaning that when interest rates rise, the price of bonds usually falls, and vice-versa. In most cases, stable economies tend to have lower interest rates and higher bond prices, ensuring safer investments, which attracts investors who seek safer assets. In this sense, interest rates play “a significant role in shaping the flow of Foreign Direct Investments” (Stefan Calimanu, 2023). “While the potential for better returns and perceived stability may be attractive, higher rates can deter FDI by increasing financial costs” (Stefan Calimanu, 2023). That is why it is important to compare and measure interest rates among EU Member States, as interest rates can showcase whether a country’s economy is stable and determine the flow of Foreign Direct Investments, which, as already mentioned, indicates a country’s external competitiveness.

Regarding tourism, “in 2023, Europe alone accounted for over half of international tourist arrivals worldwide, with inbound arrivals exceeding 700 million. Such a high influx of visitors has a significant economic impact in the whole region, with the total contribution of travel and

tourism to Europe's GDP amounting to over two trillion euros in 2023" (Statista Research Department, 2024). This suggests that the number of arrivals in all EU countries is high. Still, it is important to observe whether there is a positive correlation between eurozone membership and tourists' arrival because, similar to the relationship between interest rates and FDI, there is a positive and direct relationship between number of arrivals and FDI. This relationship, of course, is indirect because the more it is invested in tourism, the better facilities it can offer tourists, therefore attracting a higher number of arrivals into the country.

The fourth benefit I examine is eliminating transaction costs on currency exchange. With the data I took from the Bulgarian National Bank's Foreign Exchange Spot Trading against BGN for each day from June 14, 2023, until June 13, 2024, I calculated the total money the Bulgarian society lost in the exchange of euros to lev and vice versa. This is expressed by multiplying the volume of euros bought on a particular day by the fixed rate (1,95583) minus the cross-exchange rate through EUR fixing, then adding the volume of euros sold on the same day by the fixed rate minus the cross-exchange rate through EUR fixing. Then, I added the money lost together for the past 366 days starting from June 14, 2023, up to June 13, 2024.

The two main benefits that come with the elimination of money exchange are lower transaction and administrative costs and the elimination of exchange rate risks. "The overall transaction cost benefit is expected to exceed the loss of similar business revenues to banks, discussed above, giving an estimate of about 0.2 % of GDP" (Georgy Ganev, 2009). Moreover, the benefit of eliminating exchange risks, even though it is rather small, is expected to be anywhere between 0 and 0.1% of GDP, as described by Georgy Ganev. This is because currently, the exchange rate between the Bulgarian lev and euro is fixed and the risks are only potential in case a floating exchange rate is introduced. Nevertheless, eliminating money exchange is undoubtedly one of the main benefits of eurozone integration.

Data analysis

The goal of this study is to showcase the advantages of joining the eurozone for the isolated case of Bulgaria by comparing the 20 countries that have already adopted the euro currency to the other seven countries that are part of the European Union (EU) but not part of the eurozone such as Bulgaria. The other six countries apart from Bulgaria are the Czech Republic, Denmark, Hungary, Poland, Romania, and Sweden. To achieve this comparison, I conducted a longitudinal analysis. I used four models: the Unconditional Means Model (UMM), which I refer to as Model 0; the Unconditional Growth Model (UGM) as Model 1; the Growth Model + EZ (Model 2); and the Growth Interaction with EZ (Model 3). The comparison is also executed using a dummy variable for the eurozone called "EZ_rec," introduced in Model 2 and Model 3. Its value is 0 if a particular country has not joined the eurozone at a certain time and one if it has, respectively.

Model	AIC	BIC	Model 0	p-value	Model 1	p-value	Model 2	p-value
Model 0	1851.9	1862.3						
Model 1	1855.8	1876.6						
Model 2	1857.6	1881.8	2.343	0.6729	0.1887	0.664		
Model 3	1859.5	1887.3	2.3818	0.7942	0.2275	0.8925	0.0388	0.8438

Table 1: Interest Rate

Model	AIC	BIC	Model 0	p-value	Model 1	p-value	Model 2	p-value
Model 0	29250	29265						
Model 1	29103	29134						
Model 2	29075	29110	183.39	< 2.2e-16 ***	30.903	2.713e-08 ***		
Model 3	29050	29091	210.08	< 2.2e-16 ***	57.597	3.111e-13 ***	26.695	2.383e-07 ***

Table 2: Foreign Direct Investments (FDI)

Model	AIC	BIC	Model 0	p-value	Model 1	p-value	Model 2	p-value
Model 0	4435.2	4448.4						
Model 1	3638.2	3664.8						
Model 2	3640.2	3671.2	802.98	< 2.2e-16 ***	0.0534	0.8172		
Model 3	3641.9	3677.4	803.24	< 2.2e-16 ***	0.3214	0.8515	0.268	0.6047

Table 3: Tourism

	Tourism	Interest Rate	Foreign Direct Investments
Fixed effects			
Initial value (intercept)	2.2651 (4.6507)	8.17337 (3.69982)	2140.62 (3564.79)
Growth (year)	0.6058 (0.1403)	-0.11984 (0.09376)	56.14 (192.01)
Eurozone	-0.1428 (0.6177)	-0.87648 (2.01528)	26334.67 (4436.79)
Random effects			
Residual Variance	12.7668	1.442e+02	1.404e+09
Intercept Variance	548.6640	8.743e+00	1.308e+08
Year Variance	0.5074	4.629e-03	6.898e+05

Table 4: Model 2

Based on tables 1, 2, and 3, I chose model 2. Model 2 has slightly higher AIC and BIC than Models 0 and 1, but these differences are small and negligible, suggesting a comparable fit. The predictors in Model 2 have reasonable p-values, showing they contribute more meaningfully to the model than in Model 3, where the p-values suggest weaker contributions. Model 2 balances adding meaningful predictors without overcomplicating the model.

When analyzing the results for interest rate, I noticed that Model 2 has slightly worse AIC and BIC than Model 1, but the differences are negligible (within 2 units), suggesting a comparable fit. The predictors in Model 2 have reasonable p-values (e.g., eurozone $p = 0.1887$), showing they contribute more meaningfully to the model than in Model 3, where the p-values suggest weaker contributions. Model 2 balances adding meaningful predictors (like eurozone) without overcomplicating the model.

The Foreign Direct Investments (FDI) outcome emphasizes that Model 2 has a lower AIC and BIC than Models 0 and 1, suggesting a better fit. While Model 3 has the lowest AIC and BIC, the improvement is minimal compared to the added complexity. Model 2's predictors (eurozone and growth) are highly significant ($p < 2.2e-16$ for eurozone), making it a robust choice. Model 3 introduces additional complexity, and while its predictors are also significant, the improvement in fit is not substantial enough to justify the extra terms.

Regarding tourism, Model 2 and Model 3 have nearly identical AIC and BIC values, with only minor differences. This suggests the additional terms in Model 3 do not meaningfully

improve the fit. In Model 2, the eurozone predictor is insignificant ($p = 0.8172$), but the model's simplicity makes it more interpretable compared to Model 3, which adds interaction effects with no substantial improvement.

```
model_0->lmer(VariableOfInterest ~ 1 + (1|`Country Name`), DataSet ,
control=lmerControl(check.conv.singular = .makeCC(action = "ignore", tol = 1e-4)),
REML=FALSE)
Equation for Model 0
```

The fixed intercept (~ 1) represents the overall mean of the dependent variable, while the random intercept ($((1 | \text{Country Name}))$) accounts for variability in this mean across countries. The dataset (DataSet) contains the dependent and grouping variables (Country Name). The model was estimated using maximum likelihood (REML = FALSE) for model comparison, with convergence warnings for near-singular fits ignored (lmerControl). This model partitions variance into country-level differences and residual variance, forming the basis for more complex models.

```
model_1<- lmer(VariableOfInterest ~ Year + (Year|`Country Name`), DataSet,
control=lmerControl(check.conv.singular = .makeCC(action = "ignore", tol = 1e-4)),
REML=FALSE)
Equation for Model 1
```

The fixed effect of Year models the overall trend of the dependent variable over time, while the random intercept and slope ($((\text{Year} | \text{Country Name}))$) capture variability in both the baseline levels and growth rates across countries. The dataset (DataSet) includes the dependent variable, Year, and Country Name. Maximum likelihood estimation (REML = FALSE) was used for model comparison, with near-singular fit warnings ignored (lmerControl). This model allows for analyzing time-related trends while accounting for country-level differences in starting points and growth trajectories.

```
model_2 <- lmer(VariableOfInterest ~ Year + EZ_rec + (Year | `Country Name`), data = DataSet,
control = lmerControl(check.conv.singular = .makeCC(action = "ignore", tol = 1e-4)),
REML=FALSE)
Equation for Model 2
```

This model includes fixed effects for Year (capturing overall time trends) and EZ_rec (assessing the impact of eurozone membership). In contrast, the random intercept and slope ($((\text{Year} | \text{Country Name}))$) account for differences in baseline levels and time trends across countries. The dataset (DataSet) contains the dependent variable, Year, EZ_rec, and Country Name. Maximum likelihood estimation (REML = FALSE) was applied for model comparison,

with near-singular fit warnings ignored (lmerControl). This model extends Model 1 by incorporating eurozone membership as a key explanatory variable.

```
model_3 <- lmer(VariableOfInterest ~ Year * EZ_rec + (Year|`Country Name`), DataSet,
  control=lmerControl(check.conv.singular = .makeCC(action = "ignore", tol = 1e-4)),
  REML=FALSE)
Equation for Model 3
```

The fixed effects include Year (time trend), EZ_rec (eurozone membership), and their interaction (Year * EZ_rec), which examines whether the effect of time varies by eurozone membership. The random intercept and slope ((Year | Country Name)) capture variability in baseline levels and time trends across countries. The dataset (DataSet) includes the dependent variable, Year, EZ_rec, and Country Name. Maximum likelihood estimation (REML = FALSE) was used for model comparison, with convergence warnings for near-singular fits ignored (lmerControl). This model assesses the independent and combined effects of time and eurozone membership on the outcome.

It is important to mention that the variable Year was centered, and FDI and tourism were rescaled by dividing them by 1 million. Additionally, the optimizer bobyqa was used in Model 1 for FDI and tourism, in Model 2 for interest rate and tourism, and in Model 3 for tourism. Furthermore, the optimizer nloptwrap was implemented in model 2 for FDI. All this was done to deal with convergence issues and improve model performance and accuracy.

Results

I responded to my research question using four multilevel linear change models. I studied three outcomes: tourism, foreign direct investments, and interest rates.

The analysis revealed that tourism showed a significant linear growth over time across all EU countries, with a slight decrease in eurozone countries compared to non-eurozone countries, which was not statistically significant. Interest rates generally decreased across all EU countries, but these results were not statistically significant, with a more considerable decrease observed in eurozone countries, although still not significant. Foreign direct investments (FDI) increased over time for non-eurozone EU countries, while eurozone countries had significantly higher FDI. Additionally, the money lost through currency exchange between 06/14/2023 and 06/13/2024 amounted to 124,512,399.95 euros, though this figure is likely overestimated due to doubled transaction volumes.

I observed that tourism, on average, has a linear change over time. On average, the initial value for tourism in countries not part of the eurozone was 2.2651 million arrivals; however, this result is not generalizable to the population. Over the span of 56 years, tourism growth has increased on average by 0.6058 million people visiting European countries ($p < .001$). Those countries that are part of the eurozone showed a decrease in tourism by 142 thousand people compared to those not part of the eurozone; however, these results were not statistically

significant, and therefore, they cannot be generalized.

When it comes to interest rates, the initial value in countries which are not part of the eurozone averaged 8.17337. Over time it decreased with an average of -0.11984 for all EU countries. The eurozone Member States showed a decrease of -0.87648 in interest rates. The t-value for all of our estimates were not extreme enough, therefore all these results were not statistically significant.

After observing the foreign direct investments (FDI) across all EU Member States, I estimated that the FDI was initially 2.14 billion on average for countries not part of the eurozone. Over the last 56 years, FDI has increased by 56.14 million on average. However, these two estimates were not statistically significant. The eurozone Member States had a higher FDI of 26.3 billion ($p < .001$.) than the non-eurozone Member States.

Additionally, I calculated the money lost from the money exchange to be 124,512,399.95 euro for one year between 06/14/2023 and 06/13/2024. It is, however, important to mention that this amount is higher than the actual amount as the transaction volumes between banks and between banks and the Bulgarian National Bank are presented in double amounts—reflecting a purchase on one side and a sale on the other.

Discussion

Among the models tested, Model 2 was the best fit for each of these outcomes. I observed a linear increase in tourism over time across the EU, with initial tourism levels in non-eurozone countries averaging 2.2651 million arrivals. While tourism grew by an average of 0.6058 million visitors annually, eurozone members showed a slight, statistically insignificant decline of 142,000 arrivals compared to non-eurozone countries. For interest rates, non-eurozone nations had an average initial value of 8.17337, with a decrease of -0.11984 across the EU and a more substantial reduction of -0.87648 for eurozone members. Regarding FDI, non-eurozone nations saw an initial average of 2.14 billion, increasing by 56.14 million annually, while eurozone states had notably higher FDI inflows of 26.3 billion ($p < .001$.) However, most of these findings lacked statistical significance, limiting generalizability.

While the results were broadly not statistically significant, the study nonetheless suggests that euro adoption could attract higher FDI, as evidenced by the substantial increase in eurozone members. The euro simplifies cross-border investments by reducing currency-related transaction costs and exchange rate volatility, thus attracting investors. Additionally, adopting the euro might indirectly boost tourism appeal by eliminating currency exchange requirements, though this factor alone does not guarantee increased tourism flows. The observed decrease in interest rates among eurozone members could indicate a stabilizing effect of the euro, contributing to a more predictable financial environment.

The differences observed between eurozone and non-eurozone countries suggest possible economic advantages from euro adoption; however, the lack of statistical significance highlights the need for cautious interpretation. Unexpectedly, non-eurozone countries displayed greater

growth in tourism, which may reflect other influencing factors, such as regional attractions or tourism policies, rather than currency alone.

Future research should incorporate broader eurozone integration benefits, such as trade volume or inflation stability, to offer a more comprehensive view of integration impacts. A more targeted study on Bulgaria, isolating euro-related influences, could provide deeper insights, especially considering the country's unique economic conditions and potential benefits. Employing higher-quality data and advanced statistical techniques may yield more robust findings.

One of the main limitations of this study is that most findings were not statistically significant. Additionally, this research paper only considers four prospective benefits of eurozone integration, although more is to be found. Furthermore, by comparing data from eurozone Member States to data from non-eurozone countries for three of the four main variables of interest—tourism, FDI, and interest rate—the study is not entirely isolated for the specific case of Bulgaria. Also, it is important to mention that the data from the Bulgarian National Bank is not optimal as the volume of transactions between banks and between banks and the Bulgarian National Bank is presented in double the amount—as a purchase from one side and a sale from the other, meaning that the findings are higher in amount than in reality. The COVID-19 pandemic has also caused a disbalance in the data.

Despite these limitations, this study contributes to the broader dialogue on eurozone integration, providing initial insights into how adopting the euro could affect tourism, FDI, and interest rates in Bulgaria. Given the currency's impact on transaction ease and financial stability, future research and policy considerations might further explore the nuanced economic shifts that euro adoption could bring to Bulgaria's economy and regional role.

Works Cited

- European Commission. “Report on Economic and Monetary Union in the European Community.” Eurocid. Accessed September 21, 2024.
<https://eurocid.mne.gov.pt/sites/default/files/repository/paragraph/documents/27448/report-economic-and-monetary-union-european-gpeudorpdfa1bcb5689401enc001.pdf>
- European Union. “Treaty on European Union.” EUR-Lex. Accessed September 23, 2024.
<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:11992M/TXT>
- Jørgensen, Carsten. “Danish Euro Referendum 2000: Results.” Eurofound. Accessed October 5, 2024. <https://www.eurofound.europa.eu/en/resources/article/2000/denmark-votes-no-euro>
- The Directorate-General for Communication, European Commission. “Convergence Report reviews Member States' progress towards joining the euro area.” European Commission. Accessed October 5, 2024.
https://ec.europa.eu/commission/presscorner/detail/en/ip_24_3449
- ECB’s Directorate General Economics. “Convergence Report, June 2024.” European Central Bank. Accessed October 5, 2024.
<https://www.ecb.europa.eu/press/other-publications/convergence/html/ecb.cr202406~475c2172bc.en.html>
- Eglitis, Gatis, Radoslav Krastev, and Christian Weise. “Latvia: benefiting from the euro, facing new challenges.” European Commission. Accessed October 6, 2024.
https://ec.europa.eu/economy_finance/publications/country_focus/2014/pdf/cf_vol11_issue12_en.pdf
- Jurkšas, Linas. “Celebrating 20 years of the euro: the role of the euro in Lithuania.” Lietuvos Bankas. Accessed October 6, 2024.
<https://www.lb.lt/en/news/celebrating-20-years-of-the-euro-the-role-of-the-euro-in-lithuania>
- Schneider, Thomas. “Estonia and the European Debt and Economic Crisis.” Konrad Adenauer Stiftung. Accessed October 6, 2024.
https://www.kas.de/c/document_library/get_file?uuid=1da1fca8-e061-baee-fd2a-46528c2a2519&groupId=252038
- “To Euro or not to Euro: What is happening with the adoption of the common EU currency in Bulgaria.” Kapital. Accessed October 12, 2024.
https://kinsights.capital.bg/economy/2023/07/18/4508762_to_euro_or_not_to_euro_what_is_happening_with_the/
- Eurobarometer. “Introduction of the euro in the Member States not yet having adopted the common currency.” European Union. Accessed October 13, 2024.
<https://europa.eu/eurobarometer/surveys/detail/2284>
- Ganev, Georgy. “Costs and Benefits of Euro Adoption for Bulgaria.” JM-ESEE. Accessed August 17, 2024.
<https://jeanmonnetexcellence.bg/archive/assets/uploads/files/cc174%20-%20costs%20and%20benefits%20of%20euro.pdf>

Gechev, G., Beev, I., Hristozov, Y. “Expected Effects of the Euro Adoption in Bulgaria.” ResearchGate. Accessed October 19, 2024.
https://www.researchgate.net/publication/341000555_Expected_Effects_of_the_Euro_Adoption_in_Bulgaria

“Foreign Direct Investments (FDI).” European Commission. Accessed October 20, 2024.
https://single-market-scoreboard.ec.europa.eu/integration_market_openness/foreign-direct-investments-fdi_en

“Why do interest rates matter?” Federal Reserve Board. Accessed October 22, 2024.
<https://www.federalreserve.gov/faqs/why-do-interest-rates-matter.htm>

Stefan Calimanu. “How Rising Interest Rates Impact Foreign Direct Investment”. ResearchFDI. Accessed October 25, 2024.
<https://researchfdi.com/resources/articles/how-rising-interest-rates-impact-foreign-direct-investment/>

Statista Research Department. “Travel and tourism in Europe - statistics & facts”. Statista. Accessed October 26, 2024.
<https://www.statista.com/topics/3848/travel-and-tourism-in-europe/>

“Banks’ Foreign Exchange Spot Trading against BGN.” Bulgarian National Bank. Accessed July 13, 2024. <https://www.bnb.bg/Statistics/StInterbankForexMarket/index.htm>

“International tourism, number of departures.” World Bank Open Data. Accessed July 20, 2024.
<https://data.worldbank.org/indicator/ST.INT.DPRT>

“Foreign direct investments, net inflows (% of GDP).” World Bank Open Data. Accessed July 20, 2024. <https://data.worldbank.org/indicator/BX.KLT.DINV.WD.GD.ZS?locations=CB>

“Real interest rate (%) - OECD members.” World Bank Open Data. Accessed July 20, 2024.
<https://data.worldbank.org/indicator/FR.INR.RINR?locations=OE>

Comparing Classical Music Styles Using Markov Chains By Ashita Garg

Abstract

Capturing the unique musical styles of classical composers like Bach and Chopin is a challenging task due to the subjective nature of music. In this paper, we utilize Markov chains to analyze and model the stylistic elements of these composers by examining melodic and rhythmic patterns. We extract sheet music data from Bach's Preludes and Fugues and Chopin's Etudes and decompose the music into interval and rhythm frequencies using the Music21 Python library. Transition matrices are created to represent the probabilities of note and rhythm transitions, forming the basis for generating new musical excerpts through Markov chains. Our results reveal distinct differences in note and rhythm usage: Bach prefers stepwise motion and frequent use of major and perfect intervals, while Chopin displays a wider variety of rhythmic and intervallic elements. Although our generated music exhibits characteristics of each composer, it does not entirely capture their unique styles, highlighting the limitations of using only one-note dependencies. We discuss potential improvements, including extending the state space and incorporating additional musical features, to create more coherent and stylistically accurate representations.

Introduction

It takes creativity and knowledge to write sheet music. Stylistic elements that are unique to each musician will show up across their compositions. Because music is subjective, it is challenging to quantify or recreate these elements. In this paper, we attempt to capture the styles of two classical musicians, Bach and Chopin, using Markov chains. We will also look at the melodic and rhythmic occurrences in their music to see the quantitative differences in the composers' styles.

When analyzing music, some researchers took a statistical approach (Beran, Jan, and Guerino Mazzola). They translated music score data, such as metric, melodic and harmonic elements, into numerical values. This paper takes a unique approach on quantifying music as it decomposes music into hierarchical layers. Our paper takes a different approach, as our analysis focuses on the intervals and rhythm patterns of sheet music. We use the interval and rhythm frequencies from sheet music data to build Markov chains which can be used to generate random songs. Markov chains have been studied extensively in various forms. They are a random process where the next state is only dependent on the current one, commonly used to model future events. In one example, Markov chains are used as an algorithmic tool for music composition (Sericola, Bruno). It uses a suffix tree data structure, which considers time, pitch, and duration. The states of the Markov chain each represent a pattern that evaluates pitch and duration. This form of analysis considers patterns instead of individual notes, and it accounts for pitch and rhythm simultaneously. Our paper uses a different form of analysis, as we look at note and interval progressions instead of recurring patterns. Furthermore, our usage of Markov chains differs. We created two distinct chains, one for rhythm and the other for pitch. For our

experiment, we took note samples from Bach and Chopin. However, the computational pipeline is designed so that any sample of musicXML files can be inputted and used as data. To create the Markov chains, we looked at the melodic and rhythmic patterns of the sheet music and combined the chains to generate a string of notes with melodic and rhythmic components. We then used these music excerpts to see whether music generated by Markov chains can efficiently represent the style of a music sample.

Methods

Music Extraction

We analyzed the stylistic patterns of two composers: Bach and Chopin. The sheet music had to be downloaded as musicXML files, and all of these files were sourced and downloaded from MuseScore. For the Bach samples, we downloaded the Prelude and Fugue titles from the Well Tempered Clavier (titles BWV 846 to BWV 893). Additionally we downloaded Chopin's Etudes (Op. 10 and Op. 25). Some pieces were left out from Chopin's Etudes due to a difficulty in finding the xml files. There are 48 pieces from Bach's Preludes and Fugues, and there are 25 total pieces from Chopin's discography. All music is written for the piano.

We looked at two musical elements: the pitches and rhythms. The pitch of the note refers to the frequency of the sound waves - or in more simple terms, it refers to the sound that the note makes. There are 12 different notes, or 12 different pitches in the context of the piano. The rhythm refers to the duration of each note. Another musical concept that we will discuss are intervals. Intervals measure the distance between two notes and quantify it into a value. Each interval has a numerical value representing the steps, such as '2' or '7'. Attached to it will be a value of 'M', 'm', 'd', 'A', or 'P' depending on the scale in which the interval lies.

To extract the data from the musicXML files, we used the Python library Music21. All notes in the music had to be isolated, along with the subsequent note. We stored the note and rhythm values in a dataframe as a transition matrix.

Transition Matrices and Markov Chains

A transition matrix stores the probabilities of transitions between states. The rows of the matrix show the current state, and the columns show the next state. Each space of a transition matrix contains the probability of the transition from the first state to the next. These transition matrices provide the probabilities that the Markov chains require. A Markov chain is a sequence of random variables that is only generated based on the previous state in the chain. The output of the chain can be unique every time depending on the probabilities. The transition probabilities of each melodic and rhythmic note were stored in a matrix, which was then used to generate sequences of notes via Markov chains. We found these transition probabilities by creating a dataframe of each note occurrence in the sheet music followed by the note that occurs immediately after.

An example of this stochastic process is the weather. To create a transition matrix for the occurrence of a certain weather, you would only consider the weather of the previous day. If it rained, then there might be a higher chance that it will rain again the next day compared to whether it was sunny. An example is shown in figure 1.

	Sunny	Cloudy	Rainy
Sunny:	[0.7	0.2	0.1]
Cloudy:	[0.3	0.5	0.2]
Rainy:	[0.2	0.4	0.4]

Figure 1: Transition matrix representing weather events

Rhythm Transitions

There are two transition matrices for each data set. The first one is a rhythm matrix. A portion of the matrix is shown in Figure 2. This matrix contains each rhythm that occurs in the music in the rows. (i.e. 1.0, which has the value of one beat and therefore represents a quarter note). These correspond to the columns. The columns contain the rhythmic value of the note that occurs immediately after each note in the ‘Rhythm’ category. It functions as a 17x17 square matrix with the rows and columns having the same variables. The rhythms range from 0.0, which signifies a rest, and 8.0, which signifies a double whole-note.

Next Rhythm	0	1/12	0.125	1/6	0.25	1/3	0.375	0.5	0.75	1.0
Rhythm										
0.0	0.000000	0.142857	0.000000	0.000000	0.857143	0.000000	0.000000	0.000000	0.000000	0.000000
1/12	0.000000	0.866667	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.333333	0.000000
0.125	0.002567	0.000000	0.810013	0.001284	0.116816	0.000000	0.008986	0.032092	0.023107	0.001284
1/6	0.001036	0.000000	0.002073	0.964767	0.012435	0.002073	0.000000	0.015544	0.000000	0.001036
0.25	0.001573	0.000105	0.010485	0.000524	0.922093	0.000000	0.000000	0.045192	0.008388	0.007445
1/3	0.000000	0.000000	0.000000	0.022472	0.000000	0.876404	0.000000	0.000000	0.011236	0.056180
0.375	0.000000	0.000000	0.714286	0.142857	0.000000	0.000000	0.000000	0.000000	0.000000	0.142857
0.5	0.004227	0.000000	0.005917	0.008876	0.185545	0.000423	0.000000	0.718512	0.000423	0.054522
0.75	0.027211	0.074830	0.095238	0.000000	0.503401	0.000000	0.000000	0.040816	0.163265	0.074830
1.0	0.002817	0.000000	0.014085	0.005634	0.174648	0.016901	0.000000	0.405634	0.002817	0.335211
1.5	0.064935	0.038961	0.051948	0.000000	0.376623	0.000000	0.000000	0.298701	0.038961	0.012987
2.0	0.025641	0.000000	0.025641	0.000000	0.205128	0.025641	0.000000	0.205128	0.025641	0.230769

Figure 2: Transition matrix comparing rhythm occurrences

Note Transitions

The transition matrix for the note pitches proved to be more complicated, because there are multiple ways to analyze the note transitions. One method is to take the note at face value and compare how often a single pitch transitions to another. However the issue with that is that it

fails to consider the scale of the excerpt. For example, if the note ‘C’ progresses to ‘G’ at a rate of 0.7 and to ‘F’ at a rate of 0.3, we don’t gain a lot of information. Within different key signatures, each of these progressions would sound unique.

Therefore we used a second method, one that better captures the sound of the note progression: checking how often a certain interval occurs after a note position in the scale. We calculated where the note fell in the original key signature. For instance, if the note was ‘C’ in C major, it would be classified as ‘P1’. On the other hand, the columns represent how often each specific interval occurs after each note. Because of this, even though the row and column are both classified as intervals, they represent different values. The rows are each different notes with respect to key signature, and the columns are the interval jumps that happen after each note. A sample of the matrix is shown in Figure 3, created from Bach’s samples.

Interval	A1	A2	A4	A5	A6	M2	M3	M5
Note								
A1	0.000000	0.007042	0.031690	0.000000	0.000000	0.330986	0.028169	0.000000
A2	0.005236	0.000000	0.005236	0.000000	0.000000	0.366492	0.107330	0.000000
A3	0.002841	0.000000	0.008523	0.000000	0.000000	0.346591	0.056818	0.000000
A4	0.000000	0.004950	0.016502	0.000000	0.000000	0.272277	0.074257	0.000000
A5	0.005803	0.009671	0.011605	0.000000	0.000000	0.324952	0.083172	0.000000
A6	0.009091	0.001818	0.012727	0.001818	0.000000	0.283636	0.065455	0.000000
A7	0.004132	0.000000	0.000000	0.004132	0.000000	0.305785	0.103306	0.004132
A8	0.000000	0.007968	0.035857	0.000000	0.000000	0.286853	0.059761	0.000000
M1	0.000000	0.000000	0.000000	0.000000	0.000000	0.500000	0.000000	0.000000
M2	0.001199	0.003597	0.015588	0.002398	0.001199	0.347722	0.055156	0.000000
M3	0.000000	0.002663	0.013316	0.002663	0.000000	0.279627	0.126498	0.000000
M5	0.000000	0.004000	0.004000	0.000000	0.000000	0.304000	0.064000	0.000000
M6	0.001918	0.002876	0.010547	0.000000	0.000000	0.349952	0.055609	0.002876
M7	0.000000	0.009560	0.022945	0.000000	0.000000	0.269598	0.093690	0.000000
P1	0.000000	0.000000	0.007233	0.000000	0.000000	0.274864	0.097649	0.000000

Figure 3: Transition matrix comparing note occurrences

Results and Discussion

The first comparison in figure 4 examines the proportion of notes across the music samples (Bach BWV 846 to BWV 893, Chopin Op. 10 and Op. 25). The proportions are calculated based on how often a type of note takes place out of every note that the composer used. The values that are classified as ‘Note’ don’t directly refer to a note in the typical scale. They each refer to an interval. These intervals denote the distance from the tonic note of each piece’s key signature.. In other words, the graph isn’t a comparison of the exact notes that occur in the sheet music. Instead it compares the occurrence of a note that is a certain number of steps from the tonic. This method represents each composer more accurately, as it takes into account the different key signatures of every composition.

We can look at the most common types of notes used by each composer. Both Chopin and Bach tend to use more major and perfect notes in the tonic (shown by ‘M’ and ‘P’). However Bach uses a wider spread of augmented intervals, while Chopin uses diminished intervals more

frequently. This is a prominent stylistic difference in the samples. Bach uses ‘P4’ the most frequently across his Toccatas and Fugues while Chopin uses ‘M6’ and ‘P4’ at a similarly high rate.

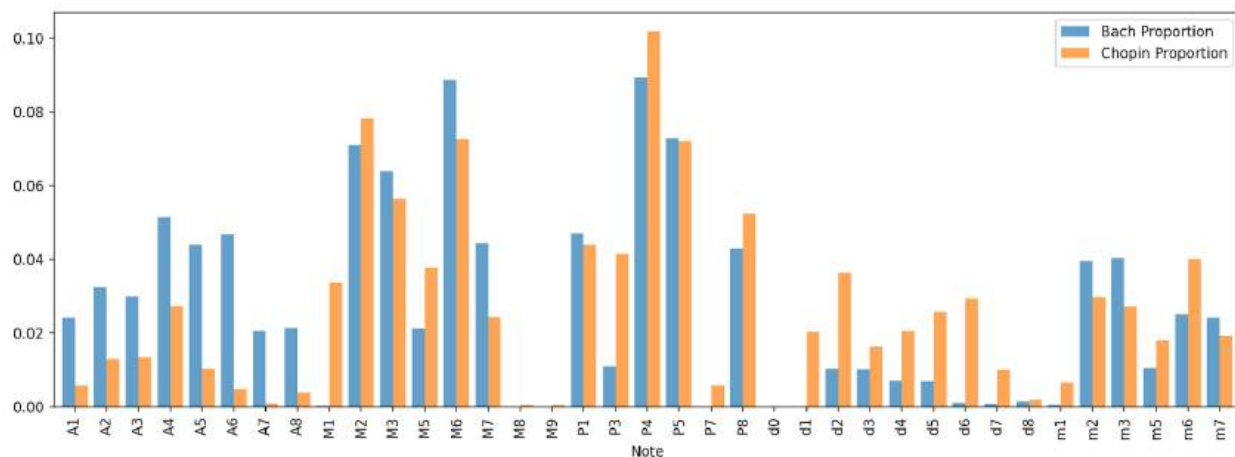


Figure 4: Frequency of different notes respective to key signature for Bach’s Preludes and Fugues and Chopin’s Etudes

Similarly to the previous figure, we created a comparison of the rhythm types in both data sets. The rhythmic values after 2.0 beats are not used frequently. However, the frequency is not necessarily zero. The most common rhythmic patterns are sixteenth notes (0.25 beats) in Bach’s Preludes and Fugues as well as Chopin’s Etudes. Bach uses these at a frequency greater than 0.6. This shows to be a key rhythmic style in the Preludes and Fugues. On the other hand, Chopin uses $\frac{1}{3}$ and $\frac{1}{6}$ beats far more frequently. This is likely due to the variances in time signatures. The beats tell us that Chopin has a lot more pieces in time signatures such as $\frac{3}{4}$ and $\frac{6}{8}$, where the distribution of a beat may be in thirds.

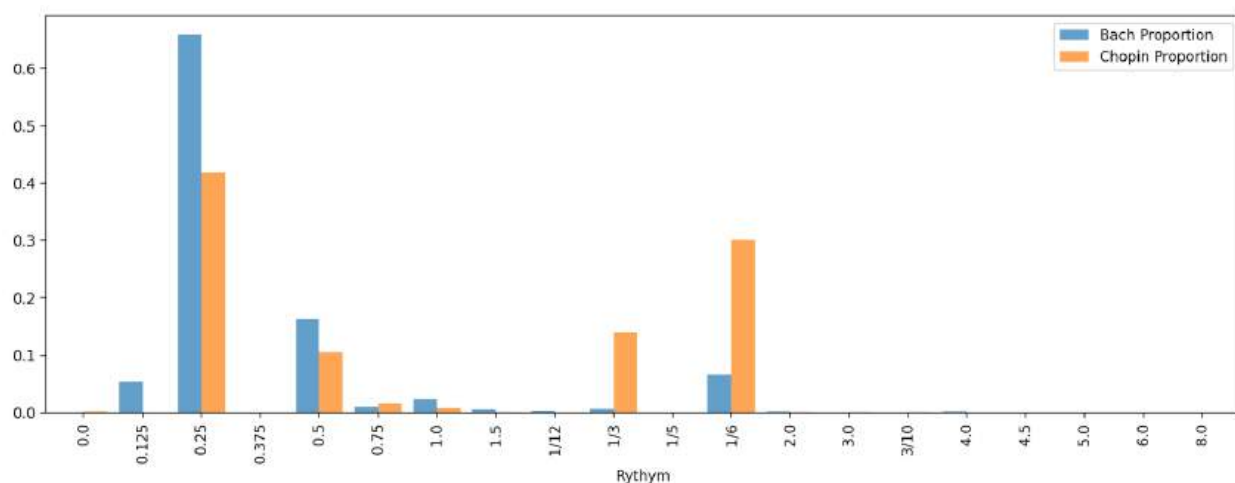


Figure 5: Frequency of different rhythms in Bach’s Preludes and Fugues and Chopin’s Etudes

The next comparison we will look at is which intervals occur the most often, shown in figure 6. These intervals differ from the ‘intervals’ in the first graph. Previously, the ‘intervals’ each corresponded to a note value that would change depending on which key signature the data is interpreted with. In this instance, the intervals can be taken at face value. Each one directly represents how frequently the composer used that spacial jump in the music. The most commonly made jump is a major 2 for both composers. However it should be noted that Bach used this interval at a rate of ~ 0.3 , while Chopin’s rate was between 0.175 to 0.200. These values represent a pattern that is repeatedly seen throughout the graphs. While Chopin has a more linear spread of intervals, Bach tends to use a couple of intervals more prominently in his Preludes and Fugues. Those intervals are major 2, minor 2, minor 3, and major 3. This data informs us that the most common steps in the music are twos and threes. Bach seldom diverges from a step motion in the pieces, rarely making big jumps between the notes. On the other hand, the frequencies in Chopin’s graph are more even, implying that he used a larger variety of intervals in his Etudes.

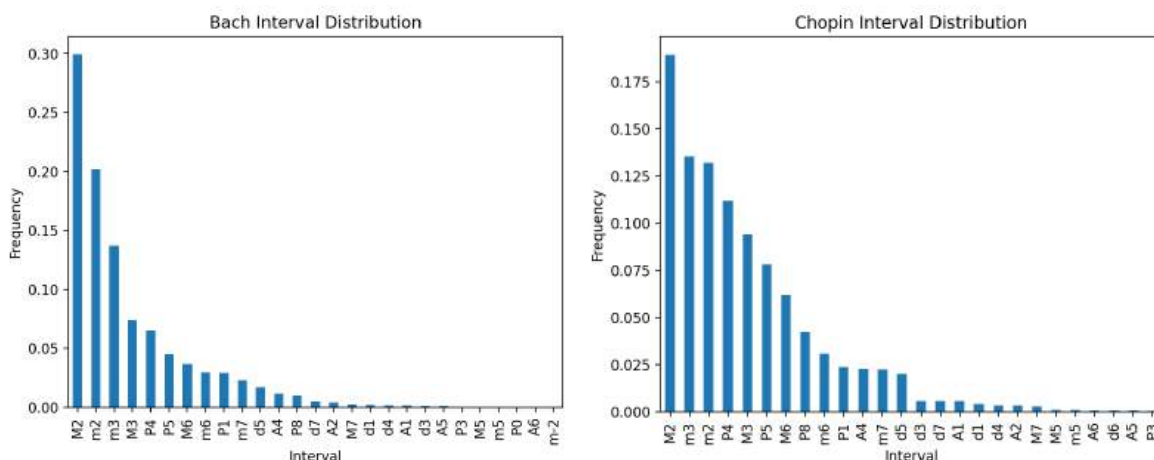


Figure 6: Frequency of interval occurrences in Bach’s Preludes and Fugues and Chopin’s Etudes. X-axis is arranged in order of decreasing frequency, which differs for each graph

Qualitative results

The result generated from the Markov chains is a one-line excerpt of music in treble clef, shown in figures 7 and 8. The parameters of key and length can be adjusted. In this instance, all generations were produced in C major. Overall, there isn’t a quantitative comparison for how drawing the data from different sources (e.g. composers) changes the output. However there are noticeable distinctions between the outputs generated for Bach and Chopin. For example, Bach has a larger variety of notes with rhythmic values of 0.25 and 0.125, while Chopin’s excerpt has a generated chain of triplets. This matches the rhythmic values shown in figure 6. Although there are only two provided music samples, an infinite amount can be generated - and each of the generations will be unique. Therefore, there isn’t one ideal excerpt of music that can be seen as a perfect representation of its composer’s style.



Figure 7: One-line excerpt generated from Markov chains of Bach's Preludes and Fugues



Figure 8: One-line excerpt generated from Markov chains of Chopin's Preludes and Fugues

Conclusion

In this experiment, we analyzed the notes and rhythms in the sheet music of various composers and analyzed the notes and rhythm. We looked at the variation in patterns between the composers. Finally, we used Markov chains to generate new random sheet music that is only influenced by the note that occurred previously. The data used to construct these Markov chains were drawn from two sources: Bach's Preludes and Fugues, and Chopin's Etudes.

There were evident patterns in the note and rhythmic progressions. Both composers have a higher frequency of major and perfect notes respective to the key signature. However, Bach tends to use augmented notes more often while Chopin tends to use diminished notes. Analysis of rhythm patterns revealed that both composers used 0.25 beat notes (eighth notes) prominently, although Bach used these far more often while Chopin had a greater variety in rhythmic types. Because of this, Bach's music often follows a stepwise progression, while Chopin's usage of $\frac{1}{3}$ and $\frac{1}{6}$ notes results in a lilting, dance-like sound, similar to the rhythm often found in waltzes and mazurkas.

Finally, we generated two excerpts of sheet music using the Markov chains. Although there are noticeable distinctions in the music, it is not sufficient to capture the individualism of each composer. This is likely because determining the notes based only on the previous note fails to account for many stylistic choices that are in music compositions. Music is subjective, and there isn't a fixed way to know if a piece 'sounds like' a certain composer. A possibility is to do a qualitative survey, where music experts can weigh in on the accuracies of the generated music. This would work by generating music from a variety of sources, and seeing whether classical musicians can identify which belongs to a given composer (such as Bach). To improve the Markov chain, we can extend the state space in the future. The next note can be chosen based on the previous two, three, or more notes that occur. It's also possible to incorporate other factors such as dynamics and other sorts of decorations. This paper uses probability to analyze and generate sheet music. It does this through dual Markov chains, one representing pitch and the other representing rhythm.

Works Cited

- Beran, Jan, and Guerino Mazzola. "Analyzing musical structure and performance---a statistical approach." *Statistical Science*, vol. 14, no. 1, 1 Feb. 1999, <https://doi.org/10.1214/ss/1009211806>.
- "Music21 Documentation¶." *Music21 Documentation - Music21 Documentation*, www.music21.org/music21docs/. Accessed 26 Nov. 2024.
- Sericola, Bruno. *Markov Chains: Theory and Applications*. Wiley-ISTE, 2013.
- Verbeurgt, Karsten, et al. "Extracting patterns in music for composition via Markov chains." *Lecture Notes in Computer Science*, 2004, pp. 1123–1132, https://doi.org/10.1007/978-3-540-24677-0_115.

Source code: <https://github.com/ashitagarg/music-generator.git>

Building the Bubble: China's Vulnerability Due to Rapid Economic Growth Fueled by Speculative Real Estate By Hans Zhu

Abstract

China's real estate sector has been a key driver of its economic rise, fueling rapid urbanization and accounting for nearly 29% of GDP by 2017 (Fung et al.). However, this growth has created systemic vulnerabilities, including high leverage, speculative overbuilding, and fiscal opacity, particularly within local government financing. These issues echo risks of the 2008 U.S. financial crisis, where excessive debt and speculative behavior led to a global economic downturn.

This paper proposes three core policy recommendations to mitigate these risks: diversifying investment opportunities to reduce household reliance on property, replacing Local Government Financing Vehicles (LGFVs) with centrally approved bonds for greater fiscal transparency, and strengthening regulatory oversight with stricter capital requirements to increase market transparency. While these measures promise long-term stability, their success hinges on careful, phased implementation to balance immediate economic needs with sustainable reform. Addressing these challenges is crucial for securing a more resilient financial future for China and avoiding a crisis with global ramifications.

Introduction

China's rapid economic growth over the past four decades has been one of the most significant global economic transformations. Between 1980 and 2020, China's GDP grew at an average annual rate of 9.5%, transforming the country from a third-world nation to the second-largest economy in the world (World Bank). With its focus on industrialization, foreign investment, and export-led growth, China's remarkable economic growth also lifted millions out of poverty, simultaneously positioning itself as a dominant player in global trade. However, beneath this impressive headline growth, structural imbalances—especially within the real estate sector—have festered, raising concerns about the sustainability of China's economic model (Fung et al.).

A central element of China's economic strategy has been its market-oriented reforms, specifically through the establishing Special Economic Zones (SEZs). First initiated in the late 1970s under Deng Xiaoping, SEZs were designed as experimental spaces to attract foreign direct investment (FDI) by offering specific capitalistic advantages not available in other parts of China. SEZs like Shenzhen, Zhuhai, and Xiamen were granted the flexibility to engage in international trade, enabling them to function as open markets with fewer state-controlled restrictions compared to other regions. Shenzhen, in particular, transformed from a small fishing village into a major industrial and financial hub, with a population exceeding 12 million by 2020 (Zeng). Across China's period of development, these zones became focal points of economic growth — especially in generating employment, promoting technological agglomeration, and bolstering export-led development.

As China's economy modernized, the construction and real estate sectors grew to become yet another engine of economic growth — by 2017, real estate and construction accounted for nearly 29% of China's GDP (Fung et al.). The types of real estate developed during this boom were varied. This included residential properties, commercial spaces, and shopping malls, as well as large-scale urban infrastructure built to accommodate the growing urban populations. Between 2011 and 2013 alone, China added approximately 4.6 billion square meters of residential real estate, primarily in major cities such as Beijing, Shanghai, and Shenzhen (Wang et al.).

The expansion of the real estate sector was closely linked to demographic changes, particularly mass urbanization. In 1980, only 19.4% of China's population lived in urban areas; by 2023, this figure had risen to 63.9% (World Bank). This migration was encouraged by government policies, including reforms to the hukou system (China's household registration system), which historically limited the movement of rural residents (Wang et al.). Between 2000 and 2015 alone, China's urban population grew by over 290 million, creating immense pressure on housing markets in the rapidly developing SEZs. In response, developers rapidly constructed residential complexes and high-rise apartment buildings to meet demand; however, these efforts provided little relief as demand continued to outpace supply and drive price increases. (Wang et al.).

This rapid expansion in the real estate market created significant vulnerabilities. The sector's reliance on debt, speculative investments, and overproduction of housing units has led to growing concerns about its sustainability. By 2021, real estate-related debt had reached 21% of China's GDP, with several major property developers — such as Evergrande — defaulting on their obligations due to liquidity issues (Fung et al.). In many cases, developers borrowed heavily to finance housing projects based on the assumption that property prices would continue to rise indefinitely. As a result, many housing developments—especially outside SEZs—were built without sufficient demand to justify their construction, resulting in the phenomenon of “ghost cities.” Such cities often consisted of entire housing complexes that remain largely unoccupied, a representation of overbuilt urban developments that were constructed based on speculative forecasts of future demand rather than immediate housing needs (Fung et al.).

Signs of fragility in the real estate sector have become increasingly apparent: Evergrande's default in 2021 sparked fears of a broader financial crisis reminiscent of the U.S. subprime mortgage crisis in 2008 (Wang et al.). High levels of leverage, weak regulatory oversight, and speculative excesses have raised concerns that China's real estate bubble could burst, with significant consequences for both the domestic and global economies. A collapse in the property market would not only lead to widespread defaults among developers but could also erode household savings, as real estate has long been regarded as a reliable and secure investment for many families.

This paper will examine the causes and effects of China's real estate bubble, contextualizing it within the broader scope of China's economic development. By comparing it to the 2008 subprime mortgage crisis in the United States, this analysis will provide insights into the risks posed by the over-leveraged and speculative nature of the real estate sector.

Furthermore, it will explore potential policy recommendations to mitigate these risks and ensure the sustainability of China's long-term economic growth.

China's Real Estate Market

In China, the real estate market operates under a state-controlled framework where the government plays a dominant role in land ownership and allocation. All urban land is state-owned, and developers lease land-use rights from local governments, typically for periods of 40 to 70 years. These land leases are a significant source of revenue for local governments, often accounting for a substantial portion of their income. As a result, local authorities have an incentive to drive up land prices, leading to higher costs for developers who rely heavily on borrowing to finance real estate projects (Liu and Xiong).

The heavy reliance on debt characterizes the market, with developers and local governments often using land and real estate assets as collateral for loans. Local Government Financing Vehicles (LGFVs) also play a significant role in funding real estate and infrastructure projects, allowing local governments to take on debt that does not appear directly on their balance sheets. These loans are often backed by future land sales or expected revenue from real estate projects, making the market highly leveraged and vulnerable to price shocks (Song and Xiong).

Literature Review

Overview

While numerous studies explore the short-term benefits of speculative growth and government interventions, long-term analyses remain limited. Although comparisons between China's real estate bubble and the 2008 Global Financial Crisis (GFC) are frequently made, they often focus on surface-level similarities without considering the institutional differences between the two events. This literature review will explore the breadth of current research, identifying key gaps in long-term perspectives and cross-event comparative analysis.

Broad Overview of China's Real Estate Market

The significant role of real estate in China's growth has been well-documented. As Liu and Xiong (2021) argue, China's hybrid economic structure—where the state wields significant control in the real estate market, while still allowing market forces to influence pricing and demand—has created a cycle of dependency. Local governments have become heavily reliant on land sales to generate operating revenue, leading to over-leverage and speculative growth in the sector (Liu and Xiong). Xiong (2023) emphasizes that real estate revenues have accounted for over 40% of local government revenues in recent years, which has created incentives to inflate land prices artificially to ensure fiscal sustainability. While these studies provide insight into the systemic risks posed by China's real estate market, they primarily focus on short-term market implications, neglecting long-term sustainability concerns. Furthermore, many point to the significant role of debt in financing real estate development in China, where local governments

and developers frequently use inflated real estate values as collateral for loans. The systemic risk lies in the fact that a downturn in real estate prices could destabilize the financial system, as developers and local governments are deeply entangled in this speculative bubble (Song and Xiong).

Short-Term Implications of Speculative Growth

In recognizing the growth of China's real estate market, much of the existing literature focuses on present impacts. Wu et al. (2015) describe the phenomenon of “ghost cities”—entire developments that remain unoccupied—as a product of overestimated demand and speculative investment, driven by local government's dependence on land sales. Moreover, Fang et al. (2016) explore how policies aimed at stabilizing the real estate market, such as easing credit to developers, have further encouraged speculative behavior. The result has been an over-accumulation of debt in the sector, with real estate-related debt accounting for approximately 25% of Chinese banks' assets (Fang et al.). Rogoff and Yang (2021) discuss the potential for declining housing demand as China's population ages and urbanization rates slow, but they do not fully explore how these long-term trends might interact with the existing debt and speculative behaviors in the real estate sector.

Comparisons to the 2008 Global Financial Crisis

There have also been many comparisons between China's real estate market and the U.S. housing bubble that precipitated the 2008 GFC. Fang et al. (2016) argue that both crises share common elements, including high leverage, speculative investment, and inflated property values. However, they point out that China's real estate sector is distinct from the U.S. market in several ways, most notably in its government-controlled credit mechanisms and high down payment requirements, which serve as buffers against a rapid collapse (Fang et al.). Wu et al. (2015) also note the parallels between the two crises but emphasize that China's state intervention in the banking and real estate sectors could prevent a sudden market crash like the one seen in the U.S. Xiong (2023) expands on this point, discussing how China's state-led financial system—where the government frequently intervenes to prevent market failures—differs fundamentally from the more capitalistic U.S. system. While China's hybrid model may reduce the likelihood of a crisis like the GFC, excessive government intervention can create moral hazard. Developers and local governments, confident that the state will bail them out in the event of a downturn, are incentivized to over-leverage. This ultimately renders government intervention as a tool to delay — not remedy — existing problems, which is an effect yet to be studied.

Summary

Existing literature highlights the short-term risks posed by speculative investment and debt accumulation in China's real estate sector. However, it often fails to consider the long-term sustainability of this growth model. Additionally, many comparisons between China's real estate bubble and the GFC remain superficial, neglecting to address the moral hazards that may shape

the outcome of a real estate crisis in China. This paper seeks to address these gaps by offering a more long-term view of the risks posed by reckless investment. Furthermore, in the context of institutional differences, this study will provide an in-depth comparison with the GFC. This is done to examine whether similar financial mechanisms—such as over-leveraging and financial contagion—could lead to a similar crisis in China’s real estate sector.

Data

China’s Current Real Estate Market

In 2023, China’s economic landscape is marked by a heavy reliance on debt, with the total debt-to-GDP ratio of 288%. This captures the immense borrowing across corporate, household, and government sectors, with the real estate market being central to debt accumulation (CEIC; IMF). The total debt amounted to approximately 285 trillion RMB (40 trillion USD), a figure that places China on par with advanced economies, despite being classified as an emerging market economy (BOFIT). One of the most concerning aspects is the significant borrowing by local governments through LGFVs, which fund a variety of real estate and infrastructure projects. By the end of 2023, local government debt alone reached 60 trillion RMB (8.5 trillion USD), equivalent to 45% of GDP. When factoring in off-balance-sheet liabilities, some estimates suggest that local government debt could exceed 110 trillion RMB (15.6 trillion USD), significantly heightening China’s overall debt burden (National Bureau of Statistics; CEIC).

At the national level, China's central government debt has also grown. By 2023, central government debt stood at 54 trillion RMB (7.7 trillion USD), or approximately 40% of GDP. This further compounds the fiscal strain on the country's broader economic health (IMF). The combination of mounting local and national debt has created vulnerabilities in the financial system, particularly as the real estate market continues to falter.

The real estate sector, which has driven China’s economic growth for decades, is now in distress. Real estate developers have amassed over 35 trillion RMB (7 trillion USD) in debt by 2023. This phenomenon is largely a result of speculative borrowing to fund housing projects, often in anticipation of endlessly rising property prices. Evergrande and Country Garden, two of China’s largest developers, have accumulated the most significant debts. Evergrande’s 2.1 trillion RMB (300 billion USD) default in 2021 sent shockwaves through the market, raising concerns about contagion risks in both domestic and global financial markets (National Bureau of Statistics).

China's banking system remains highly exposed to the troubled real estate sector. As of 2023, 25.9% of total loans in the banking sector are linked to real estate assets, including mortgages and developer loans. This high level of exposure presents a considerable risk to the financial system, especially as defaults in the real estate sector continue to rise (S&P Global). Furthermore, developers issued \$292 billion USD in corporate bonds between 2020 and 2023, but defaults on these bonds have surged to \$125 billion USD, underscoring the financial strain on the sector (IMF). Local governments issued 9.3 trillion RMB in bonds in 2023 to meet debt

repayment obligations, but this has merely delayed more significant fiscal challenges resulting from declining land sales (National Bureau of Statistics).

Real estate investment in China saw a sharp decline in 2023, falling by 9.6% year-over-year. This downturn is particularly pronounced in the housing sector, where new residential construction starts fell by 20.9%, reflecting the oversupply of housing units across the country, especially in “ghost cities” (National Bureau of Statistics). This slowdown in development is compounded by declining consumer confidence and stalled construction projects, as developers struggle to secure financing to complete ongoing projects (CEIC).

Real estate spending has also plummeted, as liquidity crises deepen across the sector. Developers have been struggling to finance existing projects, further eroding confidence in the market. As defaults rise, both real estate developers and local governments have been impacted. The international exposure to Chinese real estate bonds, many of which are issued in foreign currencies, has amplified the effects on global markets.

Local governments, which rely on land sales to generate revenue and service LGFVs, faced severe fiscal deficits as land sales revenue dropped by 19.7% in 2022. As land sales continue to decline beyond 2023, the solvency of many LGFVs is threatened, with several local governments sliding towards default. The broader fiscal implications of these revenue shortfalls are concerning, as they strain the local government’s ability to manage their debts and fund ongoing public projects (National Bureau of Statistics).

U.S. 2008 Global Financial Crisis

The U.S. 2008 Global Financial Crisis (GFC) provides a parallel case study of how debt tied to real estate can destabilize entire financial systems. A key factor in the lead-up to the crisis was the rapid rise of Mortgage-Backed Securities (MBS). By 2007, the total value of MBS in circulation had swelled to nearly \$7 trillion USD, many of which were backed by subprime mortgages—loans issued to borrowers with low credit scores. (Federal Reserve History). As housing prices surged in the early 2000s, lenders issued increasingly risky loans under the assumption that property values would rise indefinitely.

However, the housing market began to decline in 2007, leading to widespread defaults on subprime mortgages. By 2008, U.S. household debt had reached nearly 99% of GDP, with much of the debt tied to real estate loans. As housing prices fell by approximately 30% from their peak in 2006 to their trough in 2009, millions of homeowners found themselves underwater—owing more on their mortgages than their homes were worth. This rise in defaults caused the value of MBS to collapse, crippling banks and institutional investors who were heavily invested in these securities (Federal Reserve).

Banks and financial institutions were particularly vulnerable because of their extensive exposure to MBS and other real estate-backed products. Financial giants like Lehman Brothers held large amounts of MBS on their balance sheets. When the value of these securities plummeted, Lehman Brothers was forced into bankruptcy in September 2008. This marked one of the largest financial collapses in history. The U.S. government responded with the Troubled

Asset Relief Program (TARP), a \$700 billion USD intervention designed to purchase toxic assets from banks and inject capital into financial institutions to stabilize the system (IMF).

The sharp decline in housing prices led to the near collapse of the U.S. financial system. The crisis ultimately caused a significant contraction felt across Europe, Asia, and other regions. The interdependence and reliance on MBS of the U.S. financial system created a cycle of defaults and devaluations that crippled the global financial markets.

Both China's real estate crisis and the U.S. 2008 Global Financial Crisis demonstrate the significant risks posed by over-leveraged and speculative investments in real estate. China's current reliance on local government debt, coupled with speculative housing developments, mirrors many of the issues seen in the U.S. during the run-up to the 2008 crisis. The U.S. housing bubble demonstrates how a single sector's decline can shake the entire global financial system. As China continues to navigate its real estate exposure, lessons from the 2008 GFC provide insights into the systemic risks posed by leveraged speculative behavior and its broader economic threat.

Policy Recommendations

Increasing Access to Diversified Investment Opportunities

China's real estate sector has an over-reliance on property as the primary vehicle for household savings. Real estate is widely seen as an extremely safe method of investment—deeply ingrained in social norms that equate property ownership with financial security. As previously discussed, the real estate market accounted for approximately 25.9% of total loans in the Chinese banking sector as of 2023, which highlights the immense pressure placed on property investments (S&P Global). This lack of diverse investment options ultimately drives over-investment in real estate, pushing property prices above sustainable levels. To alleviate this pressure, it is crucial to provide households with access to a broader array of financial vehicles. Encouraging the development of mature financial markets, such as equities, bonds, and mutual funds, would help reduce the centrality of real estate. Diversified options can provide more stable, predictable returns, thereby decreasing the susceptibility of families' financial well-being to real-estate shocks.

Financial literacy campaigns could complement these efforts by educating citizens on the risks and benefits of different asset classes, helping overcome the deeply entrenched view of real estate as the only viable form of investment. Ultimately, this diversification strategy will address the root of speculative over-investment in property—easing systemic real estate risks and allowing households long-term economic survival.

Replacing LGFVs with Official Government Bonds

Local governments have been overly reliant on LGFVs to finance their projects, leading to unsustainable debt levels. By the end of 2023, local government debt alone reached 60 trillion RMB (8.5 trillion USD)(National Bureau of Statistics; CEIC). Originally, local governments were not allowed to issue bonds directly, which led to the creation of LGFVs as a workaround to

secure funding. As LGFVs are inherently opaque—involving multiple private banks and public entities, all intertwined—tracking the true extent of financial leverage is difficult and contributes to fiscal instability.

Replacing LGFVs with centrally approved government bonds would bring the borrowing process under stricter regulation, providing transparency and discipline to the lending sector. Government bonds approved by the central authority are easier to track and help align provincial government behavior with national economic goals. Central government approval would ensure that local governments are not over-leveraging themselves to fund unviable projects, which was a significant driver of fiscal risk under the LGFV model. Furthermore, this approach would allow investors greater insight into the real estate landscape, making their investment decisions more informed and reducing fiscal risks. A phased rollout of this policy—starting with regions most burdened by LGFV debt—would help stabilize local government finances while allowing for a smooth transition.

Strengthening Capital Adequacy Requirements and Increasing Transparency

The rapid expansion of China's real estate sector has often outpaced regulatory oversight, leading to speculative excesses and significant developer leverage. By 2021, real estate-related debt reached 21% of China's GDP, which has contributed to growing concerns about systemic risk (Fung et al.). Strengthening capital adequacy requirements would ensure developers maintain healthier balance sheets. By enforcing stricter capital requirements, there would be less reliance on borrowed capital, thereby mitigating the risk of defaults. This would help prevent over-leveraging, which has been a major factor in recent financial instability within the sector.

Transparency must also be significantly improved through the creation of a centralized, publicly accessible registry for real estate transactions and borrower financial data. Such a registry would provide consumers and investors with a clear view of the market—reducing information asymmetry and curbing speculative investments. It would also help identify areas of over-construction, such as ghost cities, which have been a persistent issue due to speculative building without sufficient demand. By focusing on these specific regulatory measures, China can create a more stable and transparent real estate sector to allow sustainable growth.

Evaluation of Policy Recommendations

The proposed policy measures target critical vulnerabilities within China's real estate sector; however, several limitations must be considered. First, promoting investment diversification away from real estate may encounter deep-rooted cultural resistance. For decades, property ownership has been synonymous with financial security in China, making it the preferred choice for household savings. Shifting this mindset requires not only introducing alternative investment vehicles such as equities, bonds, and mutual funds but also building robust regulatory frameworks to stabilize these markets. Without adequate regulatory oversight, there is a risk of replicating speculative bubbles in new asset classes, thereby transferring rather than resolving systemic financial risks.

Replacing Local Government Financing Vehicles (LGFVs) with centrally approved government bonds offers a path toward greater fiscal transparency and accountability. However, this transition is likely to face resistance from local governments that rely on LGFVs for the flexibility to secure funding for various projects. Centralized bond issuance could reduce this autonomy, potentially slowing down the execution of essential local infrastructure projects. Moreover, central oversight would not fully account for regional economic nuances, stifling local development initiatives.

Strengthening capital collateral requirements and increasing market transparency are vital to mitigating systemic risks in the real estate sector. By requiring developers to maintain healthier balance sheets, the likelihood of major defaults becomes significantly reduced. However, these measures could inadvertently limit credit availability, especially for smaller developers who may struggle to meet higher capital thresholds. This could lead to reduced market competition, fewer projects, and slower economic growth. Furthermore, building a centralized, publicly accessible registry for real estate transactions and financial data would require substantial investment in data infrastructure and robust enforcement mechanisms. Implementing these measures effectively may be a long-term endeavor, necessitating ongoing commitment from both local and central government bodies to maintain transparency and prevent speculative excesses. Despite these challenges, such measures would foster a more resilient and stable real estate market, reducing the risk of future crises and promoting sustainable long-term growth for the economy. These circumstances highlight the need for a gradual, coordinated approach to reform.

Conclusion

China's economic rise has been intricately tied to the growth of its real estate sector, which has driven urbanization, industrialization, and domestic consumption. However, the reliance on real estate as a primary engine of growth has also created significant financial vulnerabilities. High levels of debt, speculative investment practices, and the over-development of properties, particularly surrounding major urban centers, mirror issues that led to the U.S. housing market collapse in 2008. This paper highlighted the systemic risks within China's real estate sector, drawing parallels with the 2008 Global Financial Crisis to illustrate the impending broader economic instability.

To address these risks, the Chinese government should promote diversified investment options, enhance fiscal transparency with centrally approved bonds, and strengthen regulatory oversight. The effectiveness of these recommendations depends on careful implementation. A phased, region-specific rollout will reduce resistance and ensure a smoother transition. Success will depend on the government's ability to coordinate across multiple levels, ensuring that reforms do not undermine the very economic growth they aim to safeguard. Ultimately, the goal is to create a more balanced and resilient economic structure, avoiding a crisis that would reverberate beyond China's borders.

Works Cited

- BOFIT. "Chinese Debt-to-GDP Ratio Approaches 300%." Bank of Finland, 2021.
- CEIC. "China Government Debt: % of GDP." CEIC Data, 2023.
- Fang, Hanming, et al. "China's Real Estate Bubble: Lessons from the Global Financial Crisis." *Asian Economic Policy Review*, vol. 10, no. 2, 2016.
- Federal Reserve. "The U.S. Financial Crisis of 2008: Causes and Consequences." Federal Reserve History, 2021.
- Fung, Hung-Gay, Jau-Lian Jeng, and Qingfeng "Wilson" Liu. "China's Real Estate Market Development." University of Missouri-St. Louis, 2009.
- IMF. "China's Real Estate Market and Debt Structure." International Monetary Fund, 2023.
- Liu, Chang, and Wei Xiong. "China's Real Estate Market." NBER Working Paper No. 25297, National Bureau of Economic Research, 2018.
- National Bureau of Statistics. "2023 Real Estate Statistics." National Bureau of Statistics of China, 2023.
- Rogoff, Kenneth, and Yuanchen Yang. "Has China's Housing Production Peaked?" National Bureau of Economic Research, 2021.
- S&P Global. "Real Estate Sector Risks in China's Banking System." S&P Global Ratings, 2023.
- Song, Zheng, and Wei Xiong. "Risks in China's Real Estate Sector: Government Intervention and Over-Leverage." *Princeton Economic Review*, 2018, pp. 21-28.
- Wang, Xin-Rui, Eddie Chi-Man Hui, and Jiu-Xia Sun. "Population Migration, Urbanization and Housing Prices: Evidence from Cities in China." *Habitat International*, vol. 66, 2017, pp. 49-56.
- World Bank. "Urban Population (% of Total Population) - China." World Bank, 2023.
- Wu, Jing, et al. "China's Housing Boom and the Impact of the Real Estate Market on Economic Stability." *Asian Economic Policy Review*, vol. 12, no. 1, 2015.
- Xiong, Wei. "Real Estate and Financial Stability in China's Hybrid Economy." *The Arc of the Chinese Economy*, 2023.
- Zeng, Douglas Zhihua. "China's Special Economic Zones and Industrial Clusters: Success and Challenges." Lincoln Institute of Land Policy, 2012.

Bridging the Divide: Comparative Perspectives on Women's Empowerment in Developed and Developing Nations By Syed Zahir Hussain

Acknowledgments

Thank you for the guidance of Jenny Bisha, PhD, from UC Berkeley University, and Benjamin Ostrowski, PhD, mentor from Carnegie Mellon University, in developing this research paper.

Abstract

The globalization of concepts and movements has dramatically changed the understanding of gender roles in multicultural societies through the Women's Empowerment Movement. However, this critical factor of unequal empowerment for women in developing countries remains pending and has been reflected through Pakistan while creating differences between developed and developing countries. This research article undertakes a comparative analysis of the impact of women's empowerment programs in the United States and Pakistan over the last hundred years. This study evaluates critical aspects such as education, employment, political presence, and cultural domination, highlighting the differences in progress among the study countries. The review emphasizes historical incidents and socio-economic structures in the United States that have brought progressive changes in women's rights, while Pakistan has been facing cultural and systematic challenges that have not allowed for comparable change. In addition, the paper discusses the influence of globalism and social policy on the development of gender equality. Results indicate that context-sensitive strategies are necessary for spanning the empowerment gap, providing lessons in policy-making, teaching, and activism for those working hard towards women's rights causes worldwide.

Keywords Women's Empowerment, Gender Equality, Cultural Norms, Comparative Analysis, Policy Reform, Pakistan, United States

Introduction

Empowering women is fundamental to the progress of mankind, with the pace and outcome being extremely variable in developed versus developing countries. In the last century, developed nations-similar to the United States-have experienced tremendous mileage towards women's rights driven by ongoing advocacy, enabling legal frameworks, and changing cultural norms. In developing nations like Pakistan, the socio-cultural barriers are deeply entrenched and continue to hinder the movement of women's empowerment. This inequality raises the most important questions about factors that may influence the rate of progress of empowerment programs and the effectiveness of empowerment programs in different contexts. In this regard, this paper wants to answer the following question: How do socio-cultural and economic contexts influence the development of women's empowerment in developed countries and developing countries?

As noted in the global scenario, the Women's Empowerment Movement has dramatically shaped gender roles and social mores. In America, movements known as feminism which unfolded within the 19th and 20th centuries spurred on after the Seneca Falls Convention in 1848, have enabled milestones such as the passage of the 19th Amendment and the Civil Rights Act of 1964 (Alexander). These significant achievements not only conferred suffrage upon women but also established the groundwork for equity in the workplace and autonomy over reproduction. Currently, about 57% of the workforce in the United States consists of women, with nearly 40% possessing bachelor's degrees, thereby surpassing men in the realm of higher education attainment (U.S. Bureau of Labor Statistics). This noteworthy advancement signifies the collective effects of legislative changes, cultural transformations, and grassroots advocacy.

Pakistan, on the other hand, has faced some chronic systematic gender imbalances since becoming an independent country in 1947. In spite of constitutional guarantees of equality between genders promulgated by the Constitution of 1973, women in Pakistan still face some very disturbing differentials. For example, female literacy is only 65%, compared with male literacy at 79% (Aaj English TV). Additionally, the participation of women in the workforce is merely 22%, with a majority relegated to informal sectors that lack legal safeguards or opportunities for advancement (World Bank). These figures highlight the ongoing prevalence of patriarchal systems and the minimal effectiveness of empowerment programs within a societal framework resistant to cultural transformation.

It will require the study of interplay among legislative actions, cultural expectations, and grass root mobilization to understand such contradictory trends. In this regard, feminist movements within the United States challenged the dominant social structures, in which such movements have successfully influenced society to mainstream women's roles at different dimensions. Organisations like #MeToo and TIME'S UP have brought systemic issues like workplace harassment and gender-based violence into the public debate for responsibility and change (Kalachelvam). For instance, whereas in Pakistan, conservative groups oppose social initiatives such as the Aurat March on grounds of gender equity while considering themselves custodians of traditional cultural norms (Saigol). In comparison, this duality essentially points out the necessary role played by cultural openness in the success of women's empowerment initiatives.

These disparities go beyond the individual countries. Empowerment of women correlates well with positive health outcomes, economic growth, and stability in society. Empowered women play a highly important role in the economic productivity and innovation of developed countries. In the United States, for example, there has been a surge in businesses conducted by women, leading to female entrepreneurs who have been contributing \$1.8 trillion annually to the economy (U.S. Small Business Administration). By contrast, Pakistan loses an estimated \$5 billion annually due to gender gaps in economic participation (World Economic Forum). These economic ramifications emphasize the urgency of addressing empowerment disparities as a global development priority.

This paper aims to analyze these dynamics comparatively by researching the Women's Empowerment Movement in the United States and Pakistan. The study looks into key historical events, legislations' structural frameworks, and socio-cultural aspects that determine development in both countries. Based on the integration of existing literature, this study should be able to elucidate strategies that seek to bridge the gap in empowerment, specifically in developing countries. Suggest that empowering efforts should be non generic and varied and take local circumstances into mind, bearing in mind international best practices for change.

Literature Review

The Women's Empowerment Movement in the United States

The United States has been vital in the history of the Women's Empowerment Movement, from its initial steps through the First Wave of Feminism during the mid-19th century. The first wave mainly focused on filling up essential civil rights such as the right to vote and property ownership. A most significant milestone that was marked was when Elizabeth Cady Stanton promulgated the "Declaration of Sentiments" in the Seneca Falls Convention back in 1848, a revolutionary document which demanded equal rights for women (Alexander). This led to the ratification of the 19th Amendment in 1920 which granted women the vote nationally, which therefore resulted in the movement winning its first legislative victory (Jones).

However, the First Wave has been greatly criticized for its exclusionary features, particularly towards women of color and to working-class women. African American suffragists, like Ida B. Wells-Barnett, faced institutional barriers even from the feminist movement; they showed how things started to split initially in the struggle for universal gender equality (Alexander). These restrictions set the ground for future feminist movements that became focused on addressing intersectional concerns.

One of these waves was the Second Wave of feminism, lasting from the 1960s to the early 1990s and adding workplace fairness, reproductive freedom, and roles in society to the feminism focus. Betty Friedan's classic book, *The Feminine Mystique* (1963), criticized the narrow roles open to women in America after World War II, prompting middle-class women to demand greater parity at work and reproductive liberty (Jones). This period also marked a cumulative rise with legislative and other feats - like the Equal Pay Act (1963) and Title VII of the Civil Rights Act (1964). The extent of reproductive freedom guaranteed by *Roe v. Wade* of the Supreme Court in 1973 ensured that women would remain a primary beneficiary of this wave.

The political sphere had the National Women's Political Caucus that was established in 1971, but the aim was towards greater feminine participation in the government. The period also saw such big shots like Gloria Steinem who, with the founding of *Ms. Magazine* in 1972 gave feminism a prominent platform from which to use social reform for mobilization (Alexander). However, race and gender convergence became a critical focus, with Black feminists such as bell hooks offering critiques of the Second Wave for neglecting the voices of nonwhite individuals.

Gender, race, and class intersect in the Third Wave, which was a result of the movement in the 1990s. The result is that movements such as #MeToo and TIME'S UP also widened the scope of feminism in acts combating workplace harassment and sexual violence (Kalachelvam). The #MeToo movement, which started due to the revelations of #MeToo disclosures within the Hollywood entertainment industry, took the world by storm in its aftermath as millions of women shared their stories and demanded broad reforms. Nevertheless, today women are still paid 82 cents for every dollar a man earns and don't make adequate numbers within STEM fields, providing an example of what has been achieved so far (U.S. Bureau of Labor Statistics).

Another impact of the fourth wave, spearheaded by social media sites, is the promotion of gender equity and responsibility at the workplace. The United Nations' HeForShe advocacy is one of the excellent examples of the movement's continued global impact. This shows how resilient the US Women's Empowerment Movement is, for it has successfully responded to new issues in society, that it becomes a blueprint for gender campaigns around the world.

Women's Empowerment Movement in Pakistan

The Women's Empowerment Movement in Pakistan has traversed complex sociopolitical and cultural landscapes, facing such unique challenges that have arisen from traditional practices, religious interpretations, and political upheavals. The march towards gender equality in Pakistan began historically simultaneously with the country's establishment in 1947. Key figures like Fatima Jinnah advocated for women's rights, defining their importance in the nation's development. However, strong patriarchal norms limited their participation and instead made them subject to traditional roles (Saigol 12).

The first constitution, enacted in 1956, had given women double voting rights. In the same stroke of formal measures toward equality, a military coup engineered by Ayub Khan in 1958 overthrew these (Saigol 34). The 1970s would only find more reforms through the 1973 Constitution, which made particular arrangements for ensuring gender equality in the provisions of Articles 25 and 27, that Bhutto's government ushered in (Aamir). Still, social opposition to women in public roles had far higher stakes because it was buttressed by very conservative readings of religion and culture.

General Zia-ul-Haq's Islamization policies led to extreme regression of women's rights during the 1980s. The Hudood Ordinances of 1979 sanctioned severe punishments for adultery while witness account of women in rape cases was very limited and therefore subjected these women as scapegoats to the crime instead of shielding them (Saigol 67). Feminist organizations, including the Women's Action Forum (WAF), arose to protest and express their demands for gender justice (Aamir). Despite all this, Zia's regime strictly curtailed the public spheres opened for women by prohibiting women from participating in sports and regulating the media portrayal of women (Hafeez).

There was some change in the 1990s and early 2000s when Benazir Bhutto became Pakistan's first woman Prime Minister. Bhutto initiated programs such as the Lady Health Workers Program and opened women's police stations to combat gender-based violence and

bring women closer to justice (Mustafa). Nevertheless, resistance from society against the female leadership and her failure to have laws introduced by Zia nullified underscored the persistence of patriarchal structures (Aamir).

The past few years have experienced movements like the Aurat March that have brought attention to gender equality. Established in 2018, the Aurat March has grown to be a battle flag for issues in the working environment, domestic violence, and reproductive rights (Kalachelvam and Ahmed). However, powerful forces have been very vocal and combative in rejecting the Aurat March, branding it an attack on traditionalism and religious values (Hafeez).

Statistically, Pakistan trails far behind others in key gender indicators. Female literacy stands at 46% while it is a whopping 68% for its male counterpart. While female workforce participation hovers at 22%, gender-based violence pervades the country. No less than 40% of the women in Pakistan suffer physical or emotional violence at some point in their lives, and 1,000 women are killed annually in the name of honor killings (Abbasi; Mustafa). These legal developments such as the 2020 Anti-Rape Ordinance and the Protection Against Harassment of Women at the Workplace Act 2010 signify a great move forward, but difficulties in implementation emanate from weak institutional frameworks and societal opposition (Hafeez; Aamir).

For all its tenacity, the Women's Empowerment Movement in Pakistan has had to grapple with deeply rooted systemic problems. The confluence of cultural practices; weak enforcement of legal protection and limited economic opportunities underscores the imperative of all encompassing approaches towards attaining gender equity. At the root of such challenges is policymaker-people's organizations-International Partners coming together to deliver real progress.

Methodology

Using this comparative historical perspective, the analysis will track the developmental trajectory of women's empowerment in the United States and Pakistan in light of outstanding differences in their socio-cultural, political, and economic contexts. The method uses timelines from history, literature review, and data analysis that contextualize the differences by pointing out cultural, legal, and systemic aspects of how gender equity is shaped in each country. This, therefore, ensures a balanced, nuanced, and comprehensive view that fits well with the objectives of the research.

Data Collection and Sources

Most of the primary data gathered for this paper resulted from government reports, such as labor force statistics and educational attainment records. Secondary data consisted of scholarly articles, books, and historical analyses from various outlets, and international organizations such as the United Nations (UN), World Bank, and United States Agency for International Development (USAID) provided critical comparative insights. This involved exploring cross-regional thematic analyses conducted by non-governmental bodies to understand

grassroots movements and policy actions in the specific regions with regard to women's rights issues. Validity and comprehensiveness were assured since credible scholarly databases searched included JSTOR, ProQuest, and PubMed.

Comparative Framework

Based on this, a framework of systematic comparison of five major themes was formulated: education, workforce participation, political inclusion, legislative adjustment, and social change. Every one of these themes was analyzed by using such statistical metrics like literacy levels and ratios and percentage of workforce participation and representation at decisional levels. Besides, qualitative considerations like the cultural dimension, gender, and religious elements were incorporated and how they intersect with the effectiveness of a policy. This framework does allow for easy comparison of both the similarities and the differences between these two countries in terms of women's empowerment.

Chronology and Contextual Background

This historical time period analysis concentrated on the identification of crucial dates in the development of both the United States and Pakistan concerning the Women's Empowerment Movement. Important events ranging from the middle 19th century to current times, such as the ratification of the 19th Amendment in the U.S. and the institution of the 1973 Constitution in Pakistan, provided recourse points to evaluate societal changes. The researchers emphasized the assessment in terms of how these historical events influenced current efforts towards empowering women.

Limitations of the Study

Although the methodology aggregates a lot of sources and analytical frameworks, it does recognize some limitations. Data may not be uniformly available. Depending on the region, reliable and consistent information on the participation of women in Pakistan's informal sectors could be hard to come by and thus limit the scope of comparative analysis. Socio-political complexities further introduce biases in qualitative and quantitative data, especially in less developed areas. The other reality that the study recognises is that there exist significant intra-national differences due to cultural and regional diversity between different nations, which this study fails to capture.

Ethical Issues

All the sources of secondary data had been referenced carefully to maintain intellectual honesty and academic integrity. Furthermore, realistic representations of sensitive matters, such as gender-based violence and systemically systemic forms of discrimination, had aimed to ensure that research ethics were upheld also.

This methodological framework has set out an excellent foundation for looking into women's empowerment in two different socio-economic environments. Combining quantitative

data with historical and cultural analysis of the research, this book shows perspectives on what would reconcile the disparities in women's empowerment between developed and developing countries.

Comparative Analysis

Access, Participation, and Outcome Regarding Education

Education is one of the primary factors of women's empowerment, and though the difference between the US and Pakistan in this aspect is highly pronounced. Title IX of the Education Amendments passed during 1972 marked a significant landmark in the elimination of gender-based discrimination in educational environments in the US. It made it a rule that girls and women equally be provided with opportunities in institutions receiving federal funds, which naturally led to a significant rise in female enrollment in higher education. In 2020, women in the United States received about 57% of their bachelor's degrees and represented vast proportions in STEM fields, although gaps remain, especially at the leadership level in such fields.

On the contrary, Pakistan has extreme socio-cultural barriers that hinder women's access to education. Literacy among females is at 46% as compared to 68% for males; inequalities within rural areas have depicted a higher level (World Bank). Early marriages and societal norms are quite labeled as the major reasons for the high dropout rate for girls. Such initiatives, although taken forward by the Benazir Income Support Program with conditional cash transfers in support of female education, remain relatively limited in number and effectiveness given the opposition by cultural and systemic barriers.

Adaptive strategies would encompass gender-based scholarships, community involvement programs that ensure to fight cultural resistance, and policies for the establishment of women in the STEM fields also from the United States but structured according to the unique socio-cultural context of Pakistan in which grassroots advocacy and involvement of religious leaders may form integral parts of acceptance and implementation.

Workforce Participation

The labor force is another important sphere in which the United States and Pakistan differ markedly. In the U.S, the percentage of women participation is about 57%. The legal frameworks backing up this rate include the Equal Pay Act of 1963 and policies under it that outline paid family leave. Despite wage gaps that have proven to be persistent (women earn only 82 cents for every dollar made by men), there is an increase of women in leadership, as well as male-prevalent industries such as technology and finance in the U.S. (U.S. Bureau of Labor Statistics).

Indeed, in fact, women's participation in Pakistan's work-force has literally stabilized at 22% over some years now, with the majority working in agriculture and other informal sectors. Restricted mobility, low safety levels, and social expectations limit her to unpaid or low-paid skills. The Protection Against Harassment of Women at the Workplace Act of 2010 has provided

for better-safe environments, though strict implementation still remains almost neglected, particularly in rural scenarios.

Taking lessons from the U.S., showing employers' accountability, legal frameworks, and shift in sociocultural change at work, Pakistan can strengthen their hands towards women empowerment by forming women-only transport systems, vocational training for them, and promoting more micro-entrepreneurship opportunities for women.

Political Representation

Political representation is an important indicator of women's empowerment. In the United States, the year 2021 will go down in history with a record 144 women serving in Congress, including Vice President Kamala Harris. She is the first woman and first person of color to hold the position (Alexander). EMILY's List is one such organization that has been supporting women through funding to run the contests and counseling for successful electoral prospects.

Pakistan, despite ensuring the allocation of reserved seats for women within the National Assembly and Senate, has faced challenges in converting this representation into meaningful political power. Women hold 60 out of 342 seats in the National Assembly, which fulfills the 17% quota; however, this statistic signifies inadequate engagement in decision-making activities (Niazi). Cultural obstacles and patriarchal frameworks frequently marginalize women in politics, restricting their involvement to mere symbolic acts instead of actual policy formulation.

Such experiences in the United States underscore the importance of mentoring, economic support networks, and advocacy on behalf of women candidates. In the Pakistan context, grassroots involvement and channels for women leaders to interact with their constituents might even strengthen the effectiveness of reserved seats, making this potentially a tool for greater political engagement.

Societal standards and cultural constraints

Social norms and cultural perspectives play a crucial role in influencing the progression of women's empowerment across both nations. The United States has experienced pivotal movements such as #MeToo and TIME'S UP, which have illuminated pervasive problems including workplace harassment and gender-based violence. These initiatives, enhanced by social media platforms, have not only heightened public consciousness but also prompted significant reforms in corporate policies and legislation, demonstrating the efficacy of collective activism (Kalachelvam).

Part of the grassroots movements like the Aurat March in Pakistan emerge challenging deep-rooted patriarchal norms like domestic violence, reproductive rights, and workplace harassment. However, there are powerful protests by the conservative quarters who identify the activists as followers of "Western" values that shamefully belie Islam's tradition (Saigol). Such opposition demands culturally relevant strategies that amalgamate empowerment programs with religious and social values of the community.

The two streams illustrate how different experiences might be influenced by cultural openness. Although the U.S. is one of its cultural evolution destinies toward accepting gender equality, Pakistan would require an incremental, grass-roots effort fostering trust and addressing cultural concerns.

Critical Take-Aways and Lessons

In the comparative analysis, it is possible to obtain critical insights into the empowering processes for women. Developed countries, such as the United States, offer a lot of lessons on leveraging legislative reforms, grassroots movements, and cultural change toward promoting gender equality. However, the direct transfer of these tactics into the developing context often fails for different socio-cultural contexts, such as Pakistan.

Tailor Pakistan's strategy towards its challenges: For Pakistan, this axis should be of incorporating local voices in this direction, tailoring strategies to fit its needs. Let's take the example of engaging religious scholars for endorsing women's education and workforce participation against the conservative groups. Similarly, running microfinance initiatives and grassroots-level community-led education programs keep washing away resistance at the point of cultural normalcy.

This case in the U.S. and Pakistan demonstrates that no matter how vital the legal framework and policies are, sustainable empowerment comes from a conducive cultural environment. To bridge the empowerment gap, innovative context-specific strategies that balance global best practice with local realities are required.

Discussion

A comparative analysis of women's empowerment in the United States and Pakistan reveals massive inequalities, with socio-political, cultural, and economic elements playing their part. While the United States has taken gigantic strides ahead, Pakistan continues to face deeply ingrained issues that do not allow for gender equality. A proper analysis throws light on views regarding strategies and institutional barriers which affect empowerment in these disparate systems.

Socio-cultural dynamics

This theory continues to be influential in transforming American social perceptions. The international attention raised by the #MeToo movement is evident in the fact that more than 19 million tweets were posted in its first year, highlighting the prevalence of workplace harassment and gender-based violence in society (TIME, 2018). These movements challenged policies through increased accountability within the workplaces as well as the rollout of anti-harassment training within various sectors. However, the effort also ends in failure due to conservative culture and religions. For example, the Aurat March has been resisted; as of 2021, according to the report, 55 percent of the respondents rated that the movement was "Western-inspired" or not consistent with native values (Dawn, 2021).

Economic Participation

Economic empowerment is a core part of gender equality. In the United States, females make up 47% of the workforce, with over 56% working in higher-paying fields including healthcare and education (U.S. Bureau of Labor Statistics, 2023). However, discrepancies still abound, as shown through the pay gap between genders; a female earns only 82 cents for every dollar a male worker earns. In Pakistan, the situation is quite a different case, where just 22% of women are engaged in labor force activities and remain only in industries that guarantee them no formal job security or benefits (World Bank, 2023). The economic impacts of this gender inequality in Pakistan are tremendous: a closing gender gap could increase the nation's GDP by as much as 30% (IMF, 2022). However, limited mobility, insecurity, and social barriers prevent them from making the most of these chances.

It is, therefore, a changing tool.

Well, education is known generally as a prime mover of empowerment. Title IX in the United States has helped ensure that there is equal access to education; more women now graduate from college than men: 39% of adults aged 25 and older have earned a bachelor's degree or higher, and 36% of men are graduates (Pew Research Center, 2023). On the other hand, Pakistani females' literacy rate is 46%, while the males' is 71%, and the broadest ones are found in rural districts (Aaj English TV, 2024). Furthermore, UNESCO statistics show that 32% of girls in Pakistan have no enrollment at the primary level, while in the secondary level, this increases to 59% (UNESCO, 2022). All these defects call for adequate investment in school infrastructure, as well as teachers' training and public sensitization programs to overcome cultural resistance.

Political Representation and Leadership

Women's participation in politics accounts for a tremendous inequality. For instance, in the United States, women have occupied 28% of the seats in Congress with 149 individuals, with one of the most recognized examples being Vice President Kamala Harris (Pew Research Center, 2023). This results from years of grass-roots organizing and legislation that has dismantled barriers to political participation. In Pakistan, the provision of reserved seats to women in the parliament has increased their share to 20%; these positions often do not carry significant decision-making powers, though. A 2020 survey indicated that as few as 11% of Pakistani female lawmakers were involved in crucial parliamentary debates or influenced law-making (Pakistan Institute of Legislative Development and Transparency, 2020). Social mentality often reduces the capabilities of women leadership, thereby limiting their ability to bring about significant change.

Global Movements and Technology

Global mobilization and technological advancement have altered the face of advocacy campaigns. For example, #MeToo and HeForShe online social networks amplified the campaigns in the United States, engaging millions around the globe. Conversely, in Pakistan, low

digital literacy stands as a primary challenge: 14% of rural women have access to mobile internet, compared with 37% of men (GSMA, 2022). But bottom-up initiatives like the Digital Rights Foundation have been teaching over 15,000 women in digital safety and advocacy to prove that technology can bridge empowerment gaps.

Key Takeaways and Recommendations

The United States offers significant insights for developing nations such as Pakistan. The combination of legislative actions, grassroots advocacy, and the incorporation of technology has collectively propelled advancement. Nevertheless, emulating these frameworks in Pakistan necessitates a sophisticated strategy. It is crucial to incorporate cultural awareness, involve religious authorities, and implement community-focused initiatives to secure acceptance and guarantee long-term viability.

The combination of top-down policies and bottom-up social interventions shall overcome systemic barriers in Pakistan. Empowerment programs that include men as allies, promote vocational training, and create safe work environments can prove to be fruitful in these regards. Even international collaborations and funding for women-focused initiatives could further prove helpful in these regards.

This discourse asserts that though empowerment for women is possible, this will entail the continued efforts and the active formation of context-specific strategies in garnering the collective will of governments, communities, and individuals. However, it remains one of the greatest concerns regarding the world's increased gap in empowerment between developed and developing countries.

Implications and Strategies

Differences between the United States and Pakistan need wide-ranging, context-specific strategies for women's empowerment. While the U.S. is a model in legislation and cultural evolution, Pakistan's problems are unique, both socially and culturally but also economically. It has different types of obstacles needing customized solutions compatible with its traditional, religious values. The details of some crucial implications and strategic methods to fill the empowerment gap are developed in the following sections.

Policy Implications

Legal protections must be highlighted and enacted, both in the United States and in Pakistan, in advancing gender equity. In the United States, Title IX and the Equal Pay Act are two examples of statutes that have established such legally binding standards in ensuring gender equity within schools and workforces. They are examples of how consistent application can help bring incremental changes to social norms and which, overtime, can help reduce disparities between genders. Such vital enactments have been done in Pakistan, for example, Protection Against Harassment of Women at the Workplace Act 2010 and Women's Protection Bill 2006 for women rights, but inadequacies in enforcement greatly reduce their effectiveness. Pakistan

should have such independent review bodies, an accountability process for enforcement bodies which should be strong and user-friendly reporting systems ensuring that effective implementation is guaranteed. Experiences of Rwanda's post-genocide gender reforms, which imposed parliamentary representation at the legal level on grounds of gender parity, help illustrate the possibilities of focused policies leading to rapid change. Pakistan could implement similar quota systems, particularly at the local governance level, to improve women's participation and influence at the grassroots level.

Most are in the education sector, and policies that seek to expand access and participation need to be reformed to address these disparities based on gender. An all-time low of 46% of females and 68% of males have ever attained literacy; most of the required intervention is emerging. There is a need for more educational infrastructure particularly for public-private partnerships and mobile education to track geographically isolated communities. Applying the lessons learnt from Bangladesh's successful experience with the Female Secondary School Stipend Program, conditional cash transfers can be applied to Pakistan for promoting female school enrollment with corresponding reductions in dropout instances. Furthermore, female teaching competence can be built by maximizing intensive training programs and curricula that comprise gender responsive materials to assist in overcoming societal stereotypes that do not readily accept women in official occupations.

This is an area of significant improvement in both countries, although stark contrasts exist. In the U.S., for instance, legislative progress, such as the Family and Medical Leave Act and federal anti-discrimination legislation, has increased women's labor force participation to 57%. For all the lingering issues, including the pay gap, more women are gradually moving to leadership positions, especially in male-dominated fields. That of women in Pakistan is still at a nascent 22%. Most are confined to the informal sectors where job security and facilities are grossly minimum. To deal with this, one could ensure government-sponsored childcare facilities, safe transport systems, and flexible work environment policies that work for women. Some vocational training programs closely aligned with market demands—such as IT skills, textile design, and healthcare—could present suitable career paths for Pakistani women. Entrepreneurship as an option with opportunities: Microfinance programs, like Grameen Bank, provided a base by improving access to start-up capital and providing financial education for women. With the cooperation of financial institutions in designing a product especially for women, also economic barriers are reduced, and women can thereafter contribute meaningfully to the economy.

Strategic Methodologies

Strategic localization and cultural responsiveness can ensure that the women's empowerment program is implemented effectively in Pakistan. The country's cultural and religious complexities must first be recognized for entrenchment at the grassroots level. Engagement with community leaders, religious authorities, and tribal elders would facilitate contact and reduce opposition to change. Such projects as the BRAC in Bangladesh apply

empowerment strategies within conventional frameworks and prove that appropriate cultural methods can be applied to promote sustainable development. Grassroots initiatives, as those balancing women's rights with those of Islam-for example, by focusing on women's education and property rights-could further gain and change societal attitudes while continuing engagement with conservative communities.

Building structures that support women is another form of remedy. Women's resource centers in both urban and rural areas would give safe haven, legal assistance, counseling, and vocational skills. These would address the needs of women-at-risk directly. These resource centers can also act as reporting offices on cases of GBV and may link survivors to law enforcement agencies. Furthermore, advanced help lines and services for victims of harassment and domestic violence would assure an immediate intervention. In new technological offerings lie new solutions; with mobile applications that enable one to report abuse anonymously may motivate women not to fear retaliation in seeking help.

Improving digital literacy is important in addressing the significant digital gender gap that exists across Pakistan. While women are now 40% less likely to own a smartphone than men, this also prevents them from benefiting from access to telemedicine, remote education, and entrepreneurship (GSMA, 2023). Subsidized programs with free initiatives about digital literacy could help bridge such gaps, and thereby help women participate more fully in the digital economy. Adapting U.S.-based programs, such as Code.org's campaigns to promote female participation in technology, could provide young girls in Pakistan with training in STEM fields. Partnerships with schools and NGOs could make these initiatives more accessible and impactful.

Media plays a very important role in shaping people's mentality within society. Programs which appreciate the successes of Pakistani women like Malala Yousafzai and Samina Baig are excellent ways of inspiring others and making women's contribution to public life common. Media comprises television serials, social sites, and radio stations among many other mediums that can be used effectively to challenge deep-rooted stereotypes and promote gender equality. The digital platforms can also amplify the voices of grassroots movements, like the Aurat March, to much broader audiences.

The inclusion of men and boys as primary agents striving for gender equality has been fundamental. Any gender-equality campaigns with male involvement will be able to properly challenge deep-rooted patriarchy and encourage friends in the cause. Education programs aimed to enlighten men on realizing economic and social gains from women empowerment can, in the long run, end up bolstering work for gender equity. Organizations like HeForShe at the international level have shown the potential involved in engaging men into international gender advocacy. Such initiatives tailored to the socio-cultural context of Pakistan would better anchor women's empowerment in a more inclusive base so that it is responsive to all sections of society.

Main Recommendations

To move forward with the empowerment of women in Pakistan, there is a need on the part of government policymakers to adopt mechanisms that can enlist inclusiveness and

accountability. Measures need to be adopted that have inclusion for women in local governance and decision-making bodies. Strict imposition of penalties on cases of violations of workplace harassment laws and gender-based violence will help prohibit violations and ensure a safe environment for women. Provide tax incentives or recognition programs that will encourage companies to hire women, and so more people are employed in the labor force.

NGOs and civil society organizations play an important role in filling apparent gaps of government intervention. Adult women literacy programs improve gender equity in education positively by empowering more women to participate in economies and social lives. Supporting microfinance initiatives targeted at rural women can spur entrepreneurship and financial independence, thus enabling them to meaningfully contribute to their households and communities. Global partnerships are equally important because they might bring established models of women's empowerment into Pakistan and adapt relevant, effective models from other countries to the context of the region.

International collaboration offers a lot of scope for engendering gender equities in Pakistan. Those collaborations, potentially based on sharing knowledge between Pakistani organizations and American women's rights organizations, will provide some significant means of learning what works best in the policy development process as well as the implementation process. And increasing Pakistan's engagement with global empowerment efforts-ideas such as those espoused by the United Nations SDGs-will attract international attention and investment in national efforts.

Monitoring and review mechanisms ensure that the momentum created through women's empowerment is maintained. A systematic measurement of the data dispersed on key indicators, including literacy rate, participation in the labor force, as well as political participation, increases openness and helps track areas of focus. Independent review boards can even be an added tool towards further accountability, such as ensuring legislation and programs are followed expeditiously with their impact continuously measured.

Above all that relates to morality, empowering women is a strategic imperative for socio-economic development in Pakistan. Lessons from developed countries like the U.S.A. can be merged with strategies contextually appropriate and applied to Pakistan to ensure the removal of systemic barriers towards a more egalitarian society. The policy and strategy on the basis of pragmatic applications and cultural sensitivity may form the basis for a roadmap for narrowing the empowerment gap and providing true progress.

Conclusion

Women's empowerment can be seen as a trend that reflects the resilience and strength of those who have had to struggle with unequal systems for generations. The demarcation between what was achieved in the United States and what was faced in Pakistan tells us just how difficult and messy of an issue this is. While, at the same time, the United States can be regarded as the symbol of progress even in its legal reforms, societal acceptance and international movements

like #MeToo, Pakistan stands out as the challenge that many women in developing nations face every day.

This comparative analysis makes a point that women empowerment goes beyond legislative amendments or changes in policy; rather, it relates to the restructuring of societal patterns. The United States would see this as decades of activism and community-driven initiatives with a commitment to moving towards equality in the future. Pakistan's road ahead is tremendous in terms of requiring courage and innovative thinking as well as a willingness to challenge cultural norms that hostage progress. In fact, education, workforce inclusion, and political representation are not just goals; they stand as lifelines for millions of women yearning for a better tomorrow.

As readers, we think of what we can do in this global movement. Whether as an advocate, educator, or by giving voice to the marginalized, every step counts. The plight of women in Pakistan-and similarly feudal countries-is not one that is happening a thousand miles away; it beckons us to action for those who believe in the basic right to equality.

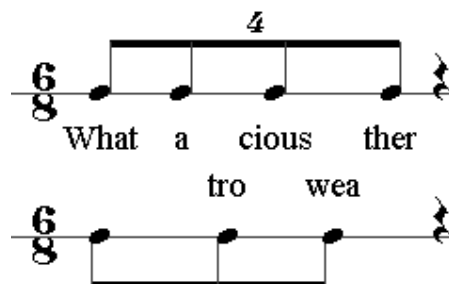
In a nutshell, women's empowerment transcends measurement or academic critique; it is the heart of a just and thriving society. The findings of this research are clear: improvement can be had, but of nothing could it be less likely to be assured. It depends on collective action, continued focus, and steered determination. Together we can learn from the success of the U.S. as well as understand and address the challenges individually to be faced by countries like Pakistan and help fashion a world where no woman is left behind.

But the fight for women's emancipation is still not yet over, truly deserving of never being done-with so many women wanting a society equal, with so many communities which happen to thrive when their women can thrive, and in a future that might finally celebrate the real potential of human beings.

Works Cited

- Aaj English TV. "Pakistan's Literacy Rate Rising Despite Decline in Education Expenditure." Aaj English TV, 2024.
<https://english.aaj.tv/news/330364928/pakistans-literacy-rate-rising-despite-decline-in-education-expenditure>.
- Aamir, A. "Feminist Movements in Pakistan: Challenges and Struggles." Capire, 2024.
<https://capiremov.org/en/analysis/feminist-movements-in-pakistan-challenges-and-struggles/>.
- Abbasi, T. R. "Branded for Life: Acid Attacks in Pakistan." Paradigm Shift, 2023.
<https://www.paradigmshift.com.pk/acid-attacks/>.
- Alexander, K. L. "Feminism: The First Wave." National Women's History Museum, 2021.
<https://www.womenshistory.org/exhibits/feminism-first-wave-0>.
- Alexander, K. L. "Feminism: The Second Wave." National Women's History Museum, 2020.
<https://www.womenshistory.org/exhibits/feminism-second-wave>.
- Alexander, K. L. "Feminism: The Third Wave." National Women's History Museum, 2020.
<https://www.womenshistory.org/exhibits/feminism-third-wave>.
- Chinni, D. "More Women Than Men Have College Degrees. That's Good News for Democrats." NBC News, 2023.
<https://www.nbcnews.com/meet-the-press/data-download/women-men-college-degrees-good-news-democrats-rcna100833#>.
- Hafeez, S. "Culture of Impunity: Is Pakistan Safe for Women?" The Express Tribune, 2023.
<https://tribune.com.pk/story/2403284/culture-of-impunity-is-pakistan-safe-for-women>.
- Kalachelvam, K., and Ahmed, I. "Aurat March: The Struggle for Law Reform and Women's Rights in Pakistan." Institute of South Asian Studies, 2021.
<https://www.isas.nus.edu.sg/papers/aurat-march-the-struggle-for-law-reform-and-womens-rights-in-pakistan/>.
- Mustafa, M. "Unfolding the Honour Killing Phenomenon in Pakistan." The Nation, 2024.
<https://www.nation.com.pk/02-Jan-2024/unfolding-the-honour-killing-phenomenon-in-pakistan>.
- Saigol, R. *Feminism and the Women's Movement in Pakistan: Actors, Debates, and Strategies*. Friedrich Ebert Stiftung, 2016.
- Smith, K. *Global Perspectives on Gender Equality*. University Press, 2021.
- UN Women. *National Report on the Status of Women in Pakistan, 2023 - A Summary*. 2023.
https://pakistan.unwomen.org/sites/default/files/2023-07/summary_-nrsw-inl_final.pdf.
- U.S. Bureau of Labor Statistics. "Labor Force Participation Rates by Gender, 2023." Bureau of Labor Statistics, 2023. <https://www.bls.gov>.
- van Baak, C., Hayes, B. E., Freilich, J. D., and Chermak, S. M. "Honor Killings in the United States from 1990 to 2021: Primary Victims and Corollary Victims." *Crime & Delinquency*, vol. 70, no. 8, 2024, pp. 2043-2072.
- Web Desk Geo News. "40% Women in Pakistan Experienced Physical or Emotional Violence." *Geo News*, 2023.
<https://www.geo.tv/latest/469238-40-women-in-pakistan-experienced-physical-or-emotional-violence>.

Math and Music: Polyrhythms By Moksh Naik



Introduction

Music; We cannot live without it. Whether we are studying, vibing with our friends, or sometimes even trying to go to sleep, we regularly find ourselves with our headphones on, listening to some beats. However, what may have slipped past your ears was the fact that all music is made in layers. Melody, bass, countermelodies, etc. These music patterns are combined to form a special, mathematical music: polyrhythmic music!

Subsection 0 Activity

If you want to observe this even further, you can do this on your own! Take out your cell phone, or any other device that plays music. Plug in your headphones, and play a song. Take into account all of the many different types of beats in the song, and read into them. Not only will this give you a glance at the concept of polyrhythmic music, but will also allow you to find a more newfound appreciation for a song.

Section 1 What are polyrhythms?

The definition of Polyhythmic music is in the name! “Poly-”, a prefix for many, and rhythmic, a series of sounds that come together in a timely fashion to produce a pattern of sounds that are neat and organized to the ears. Think about clapping your hands in sync to a 4/4 time signature, and have a friend clap in between each gap of yours. That is a basic example of a polyrhythm. Polyrhythms are easy to decipher listenin-wise, and understand what they are, but visually and mathematically will need a tad bit of breaking down, because it is a concept that cannot be understood at face value. We will learn this through some Elementary level mathematics and scientific technology.

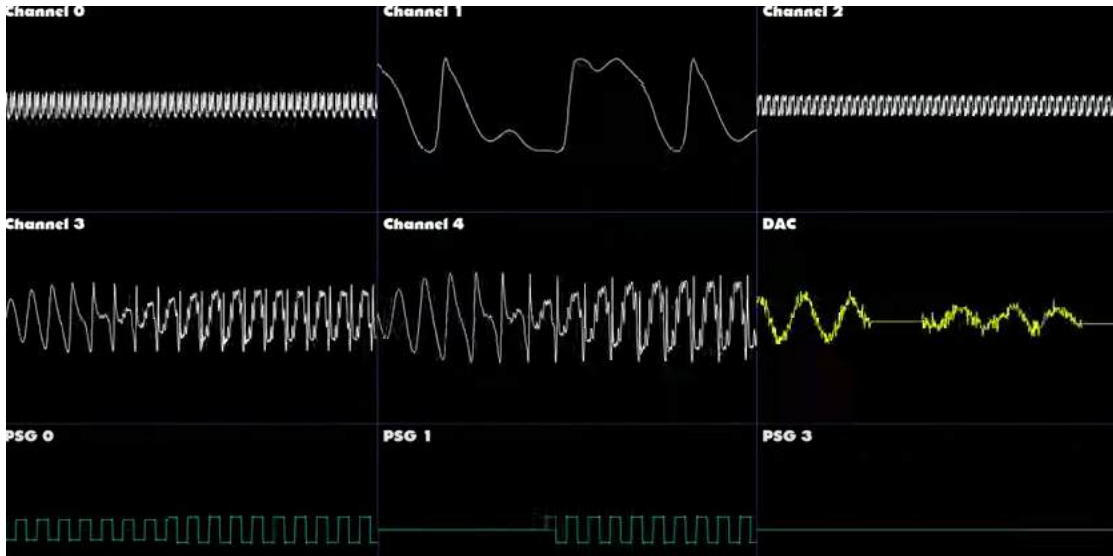
Subsection 1 Use in the real world and rhythm

Did you know that African Tribal Music and Indian Tabla music are well known as polyrhythmic music? A drummer, for example, might play 11 beats in one hand, and 12 in another. Each beat is played at a different time. However, the beats come together after a certain point. When exactly is that point? This is where mathematics comes into play. By multiplying the

The above oscilloscope is using a transducer to read an ultrasonic wave from an ultrasound machine.

Subsection 3 Oscilloscope Deconstructions of Video Game Music (Sega Genesis)

Previously we talked about how music producers use oscilloscopes in their music. Here we have an example of an oscilloscope used in video game music!



This is the Sega Genesis sound chip device analyzed through an oscilloscope. A brief history lesson for those who don't know; the Sega Genesis was a video game console released in 1989 in the United States, and the main rival to the very beloved Super Nintendo Entertainment System, or SNES for short. Both consoles had their differences, a large one being their soundchips. The SNES had better sound quality, but the Genesis was able to put out some really good tunes for what they had. Here, the channels represent different things. This image is an oscilloscope deconstruction of the loved video game soundtrack, "Hydrocity Zone Act 2 - Sonic 3 and Knuckles OST". Channels 3 and 4 are the main melody. Channel 0, PSG 0 and 1 are accompanying melodies, and in some parts of the song countermelodies. Channel 3 provides an "echo" to Channel 0. PSG 3 is a channel that makes static sounds. Channel 1 and DAC are the most important parts of the Genesis sound chip. Channel 1 is the almighty Sega Genesis bassline. DAC is the snare drums. The sound chip, despite not being as good quality as the SNES is, its songs effectively make the listener feel that if one part of the song is missing, then the whole song will not feel right at all!

Subsection 4 Genesis Soundtrack Webquest (Activity)

Look up Sega Genesis oscilloscope deconstruction videos on YouTube. Watch them and see how each channel helps each other to make really good music! Good examples include, but are not limited to: "Angel Island Zone Act 2", "Launch Base Zone Act 1" "Gene Gadget Zone

Act 1”, and “Panic Puppet Zone Act 1”. A good idea for you would be to use good quality headphones, to really bring out the life behind the sounds.

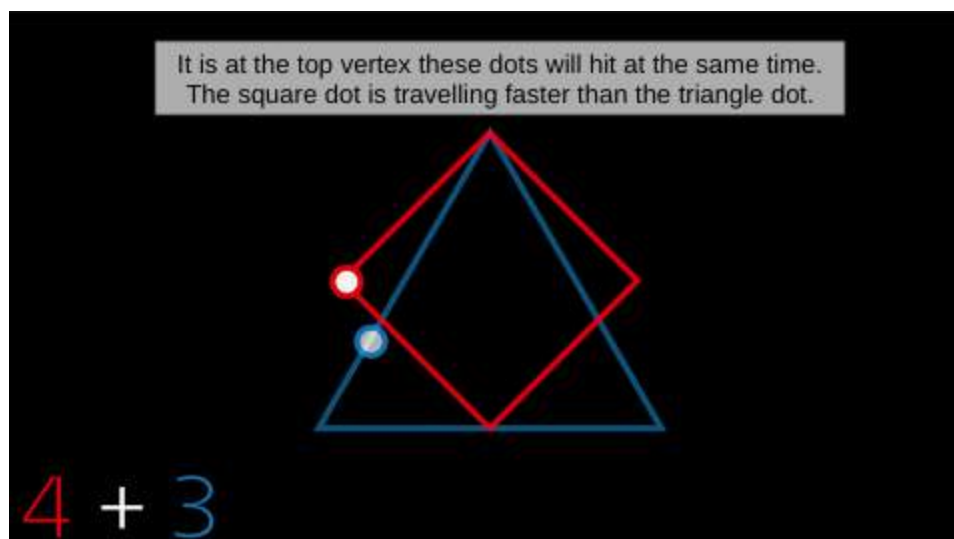
Section 3 Geometric Polyrhythms

A happy marriage for listening learners and visual learners in the comprehension of polyrhythmic music is about to follow. Geometric Polyrhythms use shapes to visualize polyrhythms. Let’s dive into the math of this.

Let’s take, for example, x as a set unit of time, which we can just make up. Then we make a ratio of $3x:4x$, where the number represents the amount of sides the shape has per x . You will get a resulting ratio of:

$$Ax : Bx$$

Next, assign a regular polygon that fits each number, in this case an equilateral triangle and a square. Now, picture a dot that moves at the speed of (ax) in each shape. What you get is that every time a dot makes a full trip around its shape, the other dot will have finished its trip as well, and have produced a nice beat while at it!



You can take this a step further by replacing the a value with other shapes, like 5 for a pentagonal 5-speed beat, 6 for a hexagonal 6-speed beat, and more!

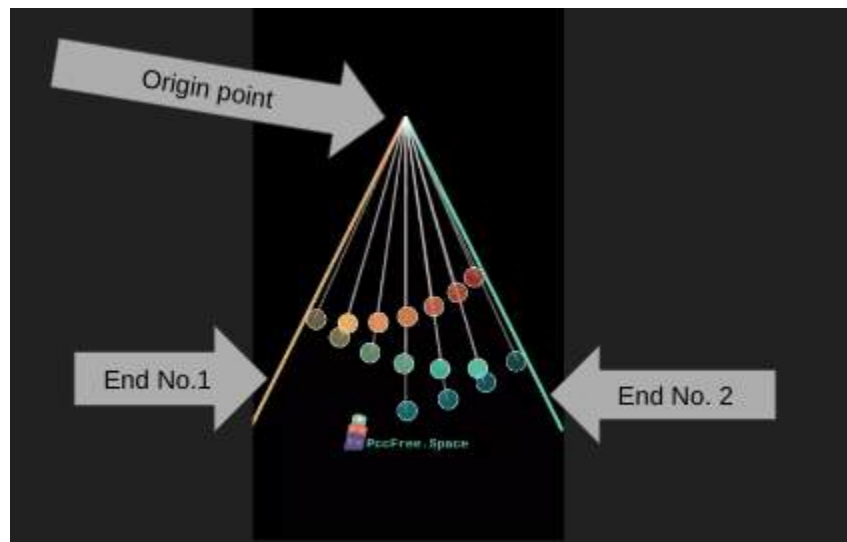
This method does not have to follow the rules of least common multiples from Section 1. In this case, the square dot will have to travel at a faster speed than the triangle dot since it has to hit more vertices. If both dots, however, DO travel at the same speed, the least common multiple rule will apply.

Since $3 \times 4 = 12$, every 12th beat the dots will hit the top vertex at the same time.

Other catchy Geometric Polyrhythms include 3:5, 3:4:5, and 2:3:4.

Subsection 5 The “Pendunome”

A “pendunome” is a made up word, but a very real concept. A “pendunome” is a phenomenon that can be described as a series of pendulums that move around a limited space, and when they come into contact with a boundary, they make a sound, swing back, and repeat the process. This is similar to what a metronome does, hence why it is called a “pendunome” since the two characteristics are present. To visualize this, think about a series of pendulums differing in string size, on the same frame. The closer you are to the point of origin, the slower you move between ends, and vice versa. When the rhythm is set in motion, you will get a very visually appealing “glissando” effect, almost to a point where your brain can automatically fill in the silence with its own glissando sound! Here is an image explaining the physics of how they work:



All the balls will start off in one direction, then slowly disperse due to their different speeds.

Conclusion

Polyrhythms are an essential basis in the music we indulge in today. Our senses enjoy listening to our favorite singer, but depending on the song, would prefer to immerse ourselves into the complexity of an instrumental, where one with enough familiarity in the world of music can understand how technical a bassline can be, where it would have to be worked on for months. The simplicity and practical genius of polyrhythms are , if anything, the most unspoken reason why we enjoy music today, because without them, let’s just say that we would be having a problem with how we listen to our music, and how we understand it as a whole.

Subsection 6 Activity

Replay the very same song you played in the first activity, but this time it is different. The reason is that you have now been filled with the knowledge and magic of polyrhythms, so that you will now have a completely new understanding of that song! Who knows? It might become your new favorite!

Works Cited

- "Poly Rhythms." YouTube, 23 Feb. 2023, youtu.be/Lxs4cvSximc?si=8_pnp2bdD_mzVeLX. Accessed 3 Mar. 2024.
- Roberts, Gareth. "Math and Music: Polyrhythmic Music." 1 Feb. 2018. College of the Holy Cross, 1 Feb. 2018, <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://mathcs.holycross.edu/~groberts/Courses/MA110/Lectures/Polyrhythms-web.pdf>. Accessed 12 Feb. 2024. Lecture.
- "Sonic 3 and Knuckles - Hydrocity Zone Act 2 - Oscilloscope Deconstruction." YouTube, 20 Dec. 2019, youtu.be/yE7HrZkHFiY?si=zbh-kQlINn91Kia7. Accessed 3 Mar. 2024.
- Tektronix. 25 Mar. 2021, www.tek.com/en/blog/what-can-an-oscilloscope-measure. Accessed 3 Mar. 2024.
- Vinci, Andrea. "What is an oscilloscope?" Tektronix, 11 May 2011, www.tek.com/en/blog/what-is-an-oscilloscope. Accessed 3 Mar. 2024.

BMI's Impact on Ability to Hold Front Lever Holds By London Bruni

Abstract

Calisthenics is divided into three disciplines: endurance, weighted, and skills. This study focuses on the front lever, which is a fundamental static hold exercise that demonstrates full body control by holding one's body parallel to the ground while hanging from a bar. Participants who actively exercise and partake in fitness were recruited and asked to submit data about their body composition, workout history, capacity to hold a front lever, and how long they could hold each front lever variation. Data was analyzed to identify trends amongst those able to hold a front lever, which revealed greater success amongst people with lower BMIs. This result will aid people in progressing their front lever holds, with implications for people interested in calisthenics, gymnastics, or even someone trying to get more fit. The emergence of this finding brings some of the first direct evaluations on factors driving front lever strength, allowing for growth and greater understanding of the discipline.

Introduction

Within the discipline of Calisthenics, there is a variety of exercises divided into three styles, including endurance, weighted, and skill calisthenics. These exercises help with different strength-improving goals. Endurance training involves performing increasing numbers of repetitions of a bodyweight exercise, weighted training involves performing repetitions with additional weight (with an external load), and skill training involves performing a gymnastic maneuver or isometric hold (Thomas et al.). This research evaluates static-hold skill training, in which an athlete completing an isometric hold will resist gravity (Azeem & Zemková).

Lever-based isometrics in calisthenics diverge from typical exercises in that body weight and body composition play a critical role in performing the calisthenics exercise. The load in the calisthenics exercise consists of one's body mass (Hass). Any weight below the shoulder will increase the force requirements to hold the lever, as one's body provides resistance in the form of a long lever arm (Hass) (Figure 1). The front lever is one of the fundamental static-hold skills in calisthenics which involves "holding [one's] body in a rigid plank... with straight arms and [one's] body perfectly parallel to the ground" (Hass).

There is a lack of research on the physiological factors impacting individuals' ability to perform static-holds. The existing research consists mostly of theoretical physics explanations. Due to the lack of research in the field, this study seeks to fill the knowledge gap about what factors may impact the ability to hold a front lever. A factor of importance in the study was Body Mass Index (BMI) score and category, a common tool used for judging whether an individual is at a healthy weight relative to their height (Cleveland Clinic Medical). BMI has been scrutinized for its limited assessment which ignores factors such as sex, age, and muscle mass (Khanna et al.); however, it is a stratifying system which allows for general assessment of future health along with a vague level of assessment of fitness, granting BMI potential in the application in sport metrics (Zierle-Ghosh & Jan). This research study analyzes body composition and workout

history as factors which impact the length of time for which a front lever is held. To do this, participants were recruited to give data on their body composition and prior workout history along with prior front lever hold times in five different variations. The collected data was separated into different tables, which diverged into two-way tables comparing one's ability to hold a front lever with factors such as BMI category.

Figure 1. Front Lever - Proper form for Front Lever Hold (*Overcoming Gravity*)



Methods

An Independent Review Board (IRB) was assembled consisting of a medical professional, science professor, and academic administrator. Participants were recruited from multiple fitness locations, including calisthenics parks, gyms, and rock climbing facilities. Participants were asked to fill out a demographic and data collection form to help establish correlations to front lever hold time. Participants were then asked to report their hold times (self-reported) in different front lever variants based on prior attempts (within the past six months). Times were analyzed for patterns and correlations with the data collected in the demographic survey. This was done by comparing the mean and median hold times for each front lever variation, along with mean and median hold times in comparison to BMI category.

Results

Seventeen individuals participated in the study, consisting of 16 males and 1 female, all of whom have some form of workout/fitness history (**Table 2**). Participants range from Healthy BMI to Overweight (18.5-29.9 BMI Score).

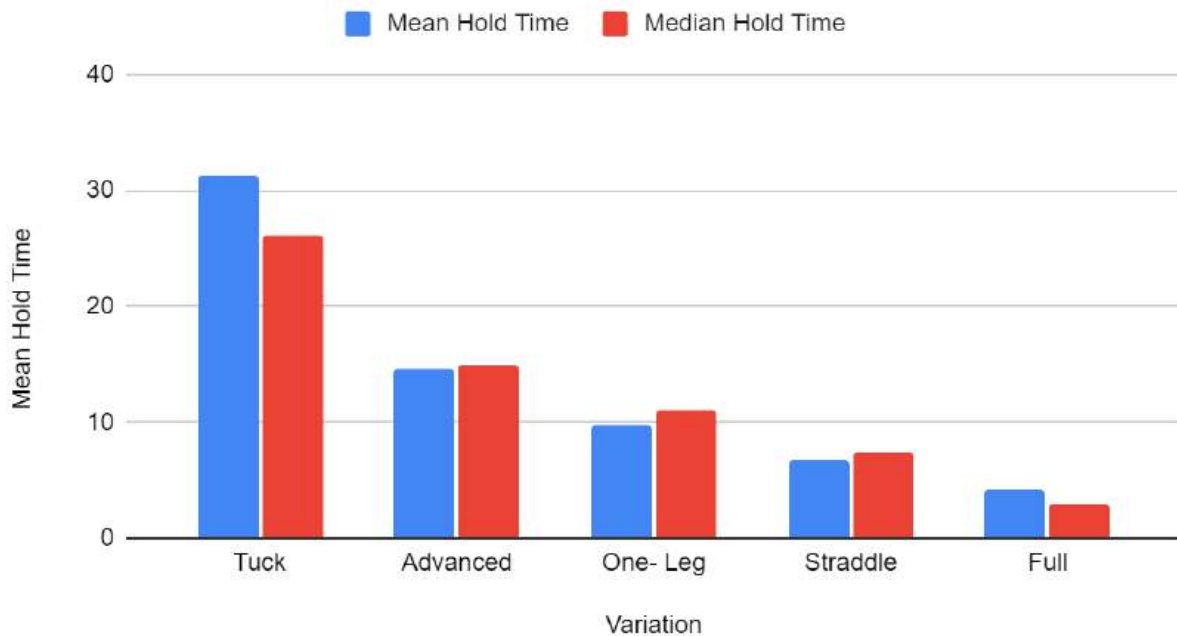
Six individuals submitted hold times. Five participants submitted tuck, one leg, advance, and full front lever hold times. Four participants submitted hold times for Straddle (**Table 1**). As participants completed each variation (Tuck Front lever to Full Front Lever), mean and median hold time per variant decreased. This is illustrated as Tuck Front lever had a mean and median hold time of 31.2 and 26 seconds, Advanced had 14.6 and 15 seconds, One-Leg had 9.8 and 11 seconds, Straddle had 6.75 and 7.5 seconds, and Full only had 4.2 and 3 seconds, respectively (**Table 1**). There was an assumed negative linear correlation across variation (**Chart 1**).

Table 1. Hold Time Data for Each Front Lever Variation.

Table 1	Tuck [s]	Adv [s]	One-leg [s]	Straddle [s]	Full [s]
P3	8	-	7	-	-
P6	0	0	0	0	0
P7	-	15	16	5	3
P10	26	12	15	12	4
P16	63	25	-	-	13
P17	59	21	11	10	1
Mean Hold Time	31.2	14.6	9.8	6.75	4.2
Median Hold Time	26	15	11	7.5	3

Chart 1. Mean Hold Time vs. Front Lever Variation

Mean/Median Hold Time vs. Variation



Of the 17 participants, the weights ranged from 53 - 92 kilograms, heights ranged from 1.55 - 1.91 meters, BMI ranged from 19-29 (Healthy to Overweight Range), and some level of workout experience ranging from on and off training for about a year to 15 years experience. (Table 2)

Table 2. Weight, Height, BMI, BMI Category, and Years of Training of Participants.

Table 2	Weight (Kg)	Height (m)	BMI	BMI Category	Training History (yrs)
P1	84	1.83	25	Overweight	4
P2	59	1.70	20	Healthy	on/off
P3	53	1.68	19	Healthy	1.5
P4	65	1.65	24	Healthy	6
P5	75	1.78	24	Healthy	3
P6	55	1.55	23	Healthy	3.5
P7	63	1.69	22	Healthy	2
P8	88	1.74	29	Overweight	10
P9	84	1.88	24	Healthy	7
P10	84	1.78	27	Overweight	2
P11	59	1.78	19	Healthy	on/off
P12	64	1.75	21	Healthy	0.5
P13	57	1.65	21	Healthy	1
P14	68	1.80	21	Healthy	3
P15	92	1.91	25	Overweight	15
P16	70	1.80	22	Healthy	9
P17	80	1.80	25	Overweight	6

Between the participants, those with the healthy BMI range (18.5-24.9) were more often able to hold a front lever compared to those with BMIs that correlated to being overweight. Out of the healthy BMI participants, 66.6% could hold the front lever, while only 60.0% of overweight BMI participants could hold the front lever (**Table 3**).

Table 3. Two-way Table of BMI category's Healthy and Overweight with self-reported ability to Hold Front Lever

Table 3	Can Hold FL	Cannot Hold FL	Total
BMI - Healthy	8	4	12
BMI - Overweight	3	2	5
Total	11	6	17

Table 4 shows that of the participants who submitted hold times, the healthy BMI individuals had a higher mean and median hold time compared to those with overweight BMIs. Healthy individuals had a mean time of 5.33 seconds while the overweight individuals had an

average of 2.5 seconds, with healthy BMI individuals holding the skill 2.83 seconds longer on average. Healthy individuals also had a median time of 3 seconds while the overweight individuals had a median of 2.5 seconds, with healthy BMI individuals holding the skill 0.5 seconds longer based on median hold time.

Table 4. Two-way Table depicting BMI category with Mean, Median, Mode Front Lever Hold Times among those who submitted times

Table 4				
Weight Category	BMI Range	Time to failure mean (s)	Time to failure median (s)	Time to failure mode (s)
Healthy	18.5-24.9	5.33	3	0,3,13
Overweight	25-29.9	2.5	2.5	1,4

Table 5 demonstrates the capacity to complete a front lever hold stratified by how many years participants have been working out.

Table 5. Two-way Table, Workout History With Ability to Hold Front Lever

Table 5	Can Hold FL	Cannot Hold FL	Total
0-2yrs Working Out	3	2	5
2-4yrs Working Out	1	3	4
4-6yrs Working Out	2	0	2
6-8yrs Working Out	1	0	1
8-10yrs Working Out	2	0	2
14-16yrs Working Out	0	1	1
On and Off	1	1	2
Total	11	6	17

There is no apparent trend that workout history increases the percentage of individuals who can hold a front lever.

Discussion

This study demonstrated that of the participants surveyed, the ability to complete a self-reported front lever hold may have been influenced by BMI, as there was an average decrease of 2.83 seconds in hold time amongst those with a higher BMI. This can be seen in

Table 1 which artificially shifts the center of mass through changing front lever progression. In progressions where body weight is shifted closer to the shoulder (fulcrum), such as the tuck front lever, hold time was typically higher relative to the full variation among those who can hold the actual front lever due to the lever arm being shorter (compact body position) leading to less force required to hold the position based on lever mechanics. This corroborates the data in **Table 3**, which shows front-lever success may be correlated with having a healthy BMI, supporting the hypothesis that having more weight or height could create a disadvantage in the lever load. **Table 4** would support the idea that those of a lower and healthier BMI have an easier time holding a front lever, as the Mean hold time in healthy BMI individuals was higher than overweight BMI individuals. A trend could not be made about workout history having an impact on the ability to hold a front lever as seen in **Table 5**, as the data seemed to lack correlation.

Calisthenics relies on a strong strength-to-weight ratio. A higher BMI, often linked to increased fat, can reduce this ratio, potentially hindering performance in calisthenics skills.

Certain pieces of data are limited as a majority of participants who consented to aid in the study did not respond to attempts of contact, resulting in less data than participants listed. Less participants with elevated BMIs submitted data, potentially biasing the results. Conclusions made from the data are limited as the number of participants was small, resulting in a potentially inaccurate representation of the population of fit individuals.

In future studies, researchers with a vast amount of participants could potentially identify the percent of fit individuals who can hold front levers, the average ratio by which changing front lever variation affects hold time, and how different styles of training (calisthenics, runner, bodybuilder, etc) can impact the ability to hold a front lever. This study lends credence to the possible findings a large sample size could reveal. In the future, more objective metrics like DEXA scans to evaluate bone density and tools to evaluate body fat compensation may reveal more data on the impacts of factors which lead to successful front-lever hold times.

Conclusion

Calisthenics is the discipline of body weight movements and within a sect of it - skills take precedent. The Front Lever is one of the notorious Calisthenics skills involving hanging from a bar and holding oneself parallel to the ground. This study intended to investigate the impact of factors such as BMI on the ability to do a front lever - finding that BMI may play a role. While there may preliminarily show an association, due to limitations in sample size more research should be done.

Acknowledgements

Thank you to Evan Case and Dr. Michael J Martinez for advising me throughout the process of writing this paper. Thank you to Aurianna Menear, Jeffery Salisbury, and Jason Jackson for participating in the Independent Review Board (IRB). Thank you as well to Kaye Regalado and the Polygence Program for writing feedback.

Works Cited

- Azeem, Kaukab, and Erika Zemková. "Effects of Isometric and Isotonic Training on Health-Related Fitness Components in Young Adults." *Applied Sciences*, vol. 12, no. 17, Aug. 2022, p. 8682. <https://doi.org/10.3390/app12178682>.
- Cleveland Clinic Medical. "Body Mass Index (BMI)." Cleveland Clinic, 28 Oct. 2024, my.clevelandclinic.org/health/articles/9464-body-mass-index-bmi.
- Hass, Tyler. "Training the Front Lever on Rings." *Crossfit Journal*, no. 63, Nov. 2007, games.crossfit.com/article/training-front-lever-rings.
- Khanna, Deepesh, et al. "Body Mass Index (BMI): A Screening Tool Analysis." *Cureus*, Feb. 2022, <https://doi.org/10.7759/cureus.22119>.
- Low, Steven. *Overcoming Gravity: A Systematic Approach to Gymnastics and Bodyweight Strength* (Second Edition). 2016.
- Thomas, E., et al. "The Effects of a Calisthenics Training Intervention on Posture, Strength and Body Composition." *Isokinetics and Exercise Science*, vol. 25, no. 3, June 2017, pp. 215–22. <https://doi.org/10.3233/ies-170001>.
- Zierle-Ghosh, Asia, and Arif Jan. "Physiology, Body Mass Index." *StatPearls - NCBI Bookshelf*, 5 Nov. 2023, www.ncbi.nlm.nih.gov/books/NBK535456.

The Neural Basis of Emotion Regulation in Major Depressive Disorder: Insights from Cognitive Behavioral Therapy and Neuroimaging Studies

By Xuehan (Helen) Ye

Abstract

Millions of people across the globe suffer from Major Depressive Disorder (MDD), which is a significant health problem. Cognitive Behavioral Therapy (CBT) influences the neural basis of Emotion Regulation (ER) in MDD patients. It is explored to make the public understand how depression is neurobiologically based as well as to develop better approaches to treating it. For this reason, a systematic literature review was conducted to summarize previous knowledge on MDD neurobiology, CBT mechanisms, and their intersection with ER. Dysregulation of neural circuits responsible for ER such as the limbic system, prefrontal cortex (PFC), and anterior cingulate cortex (ACC) is involved in MDD. The main focus of CBT involves changing dysfunctional cognitions that are implicated in depressive symptoms. Studies using neuroimaging have described changes in brain activity following CBT, especially within regions related to ER. Methodologically, investigating neural changes in MDD requires careful design and neuroimaging techniques such as fMRI, EEG, and PET. According to this analysis, the integration of ER strategies into CBT enhances treatment efficiency and neural regulation in MDD. A series of studies indicate that there is an improvement in depressive symptoms and ER after the administration of CBT

Key Words

- Major Depressive Disorder
- Emotion Regulation
- Cognitive Behavioral Therapy
- Therapeutic Strategies
- Neuroimaging

Introduction

Major Depressive Disorder (MDD) emerges as a complex and pervasive mental health challenge, casting a profound impact on the lives of millions worldwide (Bains and Abdijadid 2023, 1). According to the World Health Organization (WHO), MDD affects more than 280 million people globally (Depressive disorder (depression) 2023, 1). In the United States alone, the Centers for Disease Control and Prevention (CDC) reports that approximately 18.5% of adults experienced symptoms of depression in the year 2019 (Villarroel and Terlizzi 2020, 2). These statistics underscore the significant burden of MDD on individuals, families, and society as a whole. This debilitating condition transcends geographic and cultural boundaries, affecting diverse populations with profound and pervasive symptoms. MDD is characterized by persistent feelings of profound sadness, hopelessness, and a pervasive lack of interest or pleasure in

previously enjoyable activities (Bains and Abdijadid 2023, 1). The multifaceted nature of MDD encompasses a spectrum of symptoms that impede daily functioning, from disrupted sleep patterns and changes in appetite to difficulties in concentration and decision-making (Bains and Abdijadid 2023, 1).

Understanding the fundamental neurobiological basis of MDD is paramount for the development of effective treatment approaches. The complex connection of genetic, environmental, and neurological factors contributes to the onset of maintenance of MDD. Unveiling the neural underpinnings of this disorder holds new approaches to unlocking alternative innovative therapeutic strategies. By delving into the brain's workings in the context of MDD, researchers and clinicians can identify specific neural correlates associated with emotion regulation (ER). ER refers to the ability to effectively manage and modulate one's emotional experiences, expressions, and responses in various situations. It involves processes such as recognizing, understanding, and controlling emotions to adaptively navigate different circumstances (Schweizer, Gotlib, and Blakemore 2020, 81). In the context of MDD, ER plays a crucial role as individuals with MDD often experience significant difficulties in regulating their emotions, leading to persistent feelings of sadness, hopelessness, and despair. They may struggle to manage intense negative emotions, leading to maladaptive coping strategies or emotional avoidance, which can exacerbate depressive symptoms (Schweizer, Gotlib, and Blakemore 2020, 84). This knowledge not only enhances the public comprehension of the disorder itself but also lays the foundation for targeted interventions, such as Cognitive Behavioral Therapy (CBT), which can address the underlying neurobiological mechanisms contributing to the investigation of MDD symptoms (Gautam et al. 2020, S223).

Major strides have been made in MDD research, yet a comprehensive grasp of its neurobiological underpinnings still needs to be discovered. This paper addresses this gap, specifically focusing on the impact of CBT on the neural correlates of ER.

Understanding the neurobiological underpinnings of MDD involves recognizing the intricate interplay among brain regions and their associated neural pathways. Dysregulation within these circuits, characterized by abnormalities in neurotransmitter activity or disrupted connectivity patterns, leads to difficulties in processing and regulating emotions (Xie et al. 2024, 2). Ultimately contributing to the development and persistence of depressive symptoms, the investigation of these neural correlates gathers invaluable insights into the fundamental mechanisms underpinning MDD. It also offers potential targets for treatments aimed at restoring emotional balance in affected individuals. Regardless, despite significant advancements in understanding the neurobiology of MDD, there remain several limitations and gaps in current knowledge that underscore the need for further exploration. One of such limitations is the lack of clarity regarding the exact nature of inhibition deficits in depression and their causal relations with other observed variables. Furthermore, while previous studies have highlighted the involvement of neural circuits in cognitive inhibition and emotion regulation in depression, more research is needed to elucidate the specific mechanisms underlying these processes and their treatment implications (Jiang 2024, 2). This emphasizes the need for continued research to

examine the causal relations among the observed variables, potential interventions to improve cognitive control among depressed individuals, and the effects of training cognitive control and emotion regulation. Likewise, investigating the possibility of manipulating inhibitory control to examine its impact on the use of emotion-regulation strategies could provide valuable insights into the development of more effective therapeutic interventions for MDD. Overall, further exploration is essential for advancing our understanding of MDD and developing targeted treatments to alleviate its debilitating symptoms.

This study aims to investigate how CBT influences the neural correlates of ER in individuals with MDD. It is hypothesized that CBT interventions will exhibit discernible effects on brain structures and functions associated with ER, contributing to a nuanced understanding of the neurobiological basis of depression.

Methods

To answer this research question, a literature review was conducted. This approach involves specifically selecting and analyzing relevant research papers to synthesize existing knowledge. Here I will explore the neurobiological basis of MDD, followed by an investigation into the mechanisms and impact of CBT. The objective is to connect these bodies of research, propose avenues for further inquiry, and shed light on potential alternative research methods, aiming to synthesize key findings, identify common threads, and explore avenues for future research in the realm of emotion regulation and its significance in the context of MDD.

Search Terms and Engine

Utilizing Google Scholar, an initial exploration was conducted using the following search terms: emotional regulation, major depressive disorder, cognitive behavioural therapy, neurobiology, and therapeutic strategies.

Neurobiological Basis of Major Depressive Disorder

The complex neurobiological basis of MDD involves the dysregulation of neural circuits crucial for emotion regulation (Phillips 2003, 190). These neuroanatomical regions include the limbic system, prefrontal cortex (PFC), hippocampus, amygdala, and anterior cingulate cortex (ACC). Within these regions, abnormalities in both structure and function contribute to the emotional disturbance characteristics of MDD (Phillips 2003, 190-192).

The limbic system, particularly the amygdala, plays a crucial role in processing emotions. In the context of MDD, the amygdala often exhibits heightened activity and metabolic rates, leading to increased emotional reactivity and negative affect through its role in processing and regulating emotions. The amygdala is involved in assessing the emotional significance of stimuli and generating appropriate emotional responses. When amygdala activity increases, it becomes more sensitive and responsive to emotional cues, amplifying emotional reactions. This heightened activity can lead to an exaggerated perception of threat and danger, resulting in increased fear, anxiety, or stress responses. This hyperactivity can contribute to the

overwhelming feelings of sadness and hopelessness experienced by individuals experiencing MDD (Phillips 2003, 190). The amygdala interacts with other brain regions involved in mood regulation, such as the hippocampus, influencing the interpretation and processing of emotional information (Phillips 2003, 191-192). Conversely, the hippocampus - also part of the limbic system, playing a critical role in regulating emotions and processing memories - shows reduced volume in those with MDD (Phillips 2003, 191). This structural alteration is associated with impaired ER and memory processing, further exacerbating depressive symptoms. A smaller hippocampal volume may lead to dysfunction in the neural circuits responsible for modulating emotions, resulting in difficulties in effectively regulating mood and responding to stressors (Phillips 2003, 192). The hippocampus is involved in encoding and retrieving memories, and structural alterations in this region may disrupt the formation and consolidation of memories, exacerbating cognitive symptoms such as difficulties in concentration and recall often observed in MDD. Overall, the reduced structure of the hippocampus in individuals with MDD underscores the intricate relationship between brain morphology and the manifestation of depressive symptoms, highlighting the importance of understanding neural mechanisms in the pathology of mood disorders.

The prefrontal cortex is fundamental for cognitive control and emotion regulation, especially its dorsolateral and ventromedial regions. In MDD, abnormalities in prefrontal function are observed, including hypoactivity in the dorsolateral PFC. This reduced activity impairs the ability to regulate emotions and maintain cognitive flexibility, contributing to the persistent negative mood seen in depression (Tartt et al. 2022, 2690-2691). The abnormal activity disrupts the capacity to modulate emotional responses and adaptively shift cognitive strategies. This impairment contributes significantly to the persistent negative mood which is characteristic of depression (Tartt et al. 2022, 2690-2691). The underlying mechanisms of this hypoactivity involve dysregulation in neurotransmitter systems, particularly serotonin, dopamine, and norepinephrine, which play integral roles in mood regulation and cognitive functioning (Rădulescu et al. 2021, 3-5). The structural and functional alterations in the dorsolateral PFC, such as reduced gray matter volume and abnormal connectivity patterns, further contribute to the diminished functionality in individuals with MDD (Rădulescu et al. 2021, 2). Environmental factors, genetic predispositions, and stress also play roles in shaping these neurobiological abnormalities (Cui et al. 2024, 3).

In addition, dysregulation of the anterior cingulate cortex (ACC) – a critical area involved in processing and regulating emotions – in individuals with MDD manifests as difficulties in effectively regulating emotions and controlling emotional responses further exacerbating emotional disturbances in MDD. When dysregulation of the ACC occurs, it disrupts its normal functioning. This leads to difficulties in effectively regulating emotions, exacerbating the emotional disturbances characteristics of MDD. The ACC plays a crucial role in modulating emotional responses to external stimuli and regulating emotional experiences. It helps assess the emotional significance of events and initiate appropriate emotional responses. However, in MDD, abnormalities in the ACC impair its ability to perform these functions optimally.

Furthermore, the dysregulation of the ACC interacts with other brain regions implicated in the ER, such as the amygdala and prefrontal cortex, creating a complex network of dysfunction. The ACC has extensive connections with these regions, and alterations in its function can disrupt communication within this network, contributing to the persistence and severity of emotional disturbances in MDD (Apps, Rushworth, and Chang 2016, 1). One important subregion of the ACC is the subgenual anterior cingulate cortex (sgACC), which exhibits hyperactivity and elevated metabolic rates in comparison to individuals without MDD. This abnormal metabolic activity is indicative of heightened neural processing or dysfunction within the sgACC, potentially contributing to the emotional disturbances characteristic of MDD. Regardless, this suggests a potential target for therapeutic interventions (Pizzagalli and Roberts 2022, 227). Understanding the abnormal metabolic activity in the sgACC provides crucial insights into the neurobiology of MDD. By identifying heightened neural processing or dysfunction within the sgACC, researchers can pinpoint specific neural mechanisms underlying MDD. This knowledge not only enhances our understanding of the disorder but also opens up opportunities for developing targeted therapeutic interventions. By targeting the sgACC and its associated pathways, treatments could potentially modulate emotional processing and alleviate symptoms of depression, offering new hope for individuals suffering from MDD. Therefore, investigating and addressing abnormalities in the sgACC holds promise for advancing our understanding and treatment for MDD, ultimately improving the lives of those affected by this debilitating condition. Furthermore, abnormalities in blood flow patterns within these regions have been observed in MDD, reflecting disruption in neural communication and functioning (Singh and Gotlib 2014, 61).

Cognitive Behavioral Therapy: Mechanisms and Impact

CBT stands as a cornerstone in psychotherapeutic intervention, acclaimed for its efficacy in addressing a myriad of psychiatric disorders (Nakao et al. 2021, 1-2). CBT is a structured and evidence-based psychotherapeutic approach that aims to identify and modify dysfunctional patterns of thinking, behaviour, and emotion. It operates on the premise that our thoughts, feelings, and behaviours are interconnected, and by targeting and changing negative thought patterns and maladaptive behaviours, individuals can experience improvements in mood and overall functioning (Nakao et al. 2021, 1-2). In CBT for MDD, therapists work collaboratively with clients to identify and challenge negative automatic thoughts and cognitive distortions, such as catastrophizing and self-blame, which contribute to the maintenance of depressive symptoms (Nakao et al. 2021, 1-2). Stemming from the premise that maladaptive thoughts contribute to emotional distress and dysfunctional behaviours, CBT targets cognitive processes to foster enduring change (Nakao et al. 2021, 1-2). Notably, a seminal study by Beck and Colleagues (1979) illustrated that individuals receiving CBT exhibited significant reductions in depressive symptoms by challenging distorted thought patterns. This cognitive reframing not only mitigated negative effects but also fostered adaptive coping strategies. Additionally, neuroimaging studies have unveiled the neural underpinnings of CBT-induced ER. For example, Goldin et al. (2013)

explored that CBT interventions modulate activity in prefrontal cortical regions implicated in cognitive control while attenuating amygdalar reactivity to emotional stimuli. This neural plasticity denotes the capacity of CBT to reshape neural circuitry underlying emotional processing, thereby bolstering self-regulation mechanisms. Consequently, CBT can not only lessen disruptive recurring behaviours but can also heighten self-awareness and attention, fostering adaptive responses to emotional stimuli (Nakao et al. 2021, 1-3). These findings collectively illuminate the intricate interplay between cognitive restructuring, ER, and neural plasticity in the therapeutic efficacy of CBT, indicating its significance in alleviating psychological distress and promoting well-being.

Methodology for Investigating Neural Correlates

Investigating neural changes in individuals with MDD, particularly in the context of CBT interventions for MDD, presents several challenges. It requires meticulous methodologies to differentiate between neural changes caused by MDD itself and those resulting from the effects of CBT treatment on MDD. Specifically, this refers to studying the neural correlates of MDD before and after undergoing CBT interventions. By examining how the brain's structure and function may change in response to therapy, researchers aim to better understand the mechanisms underlying the effectiveness of CBT and identify potential biomarkers or predictors of treatment response. However, conducting such studies requires careful design, including the selection of appropriate control groups, standardized assessment tools, and rigorous data analysis techniques, to ensure the validity and reliability of the findings. Moreover, factors such as the heterogeneity of MDD symptoms, individual differences in treatment response, and ethical considerations add further complexity to the research.

One prominent challenge involves the inherent complexity of MDD and multifaceted symptomatology, making it challenging to pinpoint specific neural mechanisms underlying therapeutic interventions. Unlike some other psychiatric conditions that may present with more clearly delineated symptom profiles, MDD encompasses a wide range of symptoms, including persistent feelings of sadness, hopelessness, changes in appetite or sleep patterns, fatigue, and difficulties in concentration and decision-making (Bains and Abdijadid 2023, 1). This diversity of symptoms can vary greatly among individuals and may overlap with symptoms of other psychiatric disorders, complicating diagnosis and treatment. Additionally, MDD often co-occurs with other mental health conditions such as anxiety disorders or substance abuse disorders, further adding to its complexity (Bains and Abdijadid 2023, 1). MDD is influenced by a combination of genetic, environmental, and psychosocial factors, making it challenging to isolate and study specific neural mechanisms underlying the disorder (Bains and Abdijadid 2023, 1). These factors contribute to the difficulty in designing studies that can accurately capture the heterogeneous nature of MDD and elucidate its underlying neurobiology, hindering efforts to develop targeted and effective therapeutic interventions. Methodologically, researchers employ various neuroimaging techniques such as functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG), and Positron Emission Tomography (PET) to investigate neural

effects associated with CBT interventions for MDD (Yen et al. 2023, 1). fMRI can measure changes in the blood flow and neural activity in the brain. This technique allows for the assessment of changes in specific brain regions of interest. This can provide insights into the neural circuits involved in depressive symptomatology and their modulation through CBT (Kotoula et al. 2023, 1).

EEG assesses neural oscillations and event-related potentials, offering temporal dynamics of activities occurring within the brain during CBT sessions and potential biomarkers for treatment response (Rangaswamy and Porjesz 2008, 238). EEG measures brain changes by detecting the electrical activity generated by neurons in the brain. During EEG recording, electrodes are placed on the scalp to detect these electrical signals. As neurons communicate with each other, they produce rhythmic electrical oscillations, known as neural oscillations, which vary in frequency and amplitude depending on the brain state and activity (Bains and Abdijadid 2023, 1). Additionally, when the brain responds to stimuli or performs cognitive tasks, it generates event-related potentials (ERPs), which are transient changes in electrical activity that reflect specific cognitive processes. By analyzing the patterns and characteristics of neural oscillations and ERPs recorded via EEG, researchers can gain insights into the temporal dynamics of brain activity during CBT sessions for MDD (Liu et al. 2024, 2). These EEG-derived measures provide valuable information about how the brain responds and adapts to therapeutic interventions, potentially serving as biomarkers for treatment response.

PET imaging provides information about neurotransmitter systems and glucose activity, aiding in understanding neurochemical changes associated with CBT interventions for MDD (Finnema et al. 2015, 4144-4145). PET works by detecting positron-emitting radioactive tracers that are introduced into the body and bind to the specific molecules of interest, such as neurotransmitters or glucose. These tracers emit positrons, which collide with electrons in the surrounding tissue, resulting in the emission of gamma rays. PET scanners detect these gamma rays and create three-dimensional images that reflect the distribution and concentration of the tracer within the brain. By using tracers that bind to neurotransmitter receptors or metabolic markers like glucose, PET imaging can provide information about neurotransmitter systems' activity levels and regional brain metabolism. Consequently, PET enables researchers to observe changes in neurochemical function and metabolic activity associated with CBT interventions, shedding light on the neural mechanisms underlying therapeutic outcomes in MDD (Poli et al. 2022, 7).

Experimental designs often involve longitudinal studies assessing neural changes before and after CBT sessions, randomized controlled trials comparing neural responses between CBT and control groups, and correlational analyses examining the relationship between neural activity and clinical outcomes.

CBT Impact on Neural Correlates in MDD

Berking and colleagues (2013) conducted a randomized controlled trial that investigated the impact of integrating emotion regulation training (ERT) into routine CBT for MDD. This

study involved 432 individuals with MDD, with half receiving routine CBT and the other half receiving CBT enriched with ERT. The results indicated that participants in the CBT-ERT group demonstrated a significantly greater reduction in depressive symptoms, higher response rates, and higher remission rates compared to those in the routine CBT group. Additionally, the CBT-ERT group showed a significantly greater reduction in negative affect and an increase in well-being and emotion regulation skills relevant to mental health. The findings suggest that integrating strategies targeting emotion regulation skills can enhance the efficacy of CBT for MDD. However, the study faced challenges related to the allocation of participants to the CBT-ERT condition due to limited treatment sessions to assess quantitative data on treatment protocols. However, therapists in both the routine CBT and CBT-ERT conditions were highly trained and supervised, and there were significant group differences in emotion regulation skills, suggesting specific effects of ERT on the outcome (Berking et al. 2013, 241-243). Relatedly, these findings suggest that integrating strategies targeting ER skills, such as those used in ERT, can lead to improvements in depressive symptoms and ER in individuals with MDD. This indicates that CBT, particularly when incorporated with ERT, may positively impact the neural correlates of emotion regulation in individuals with MDD. Further research specifically focusing on the neural correlates of ER in the context of CBT for MDD would be valuable to provide a more comprehensive understanding of its impact.

Another research study, Kleefa (2022), focused on the neural correlates of ER in individuals with remitted recurrent Major Depressive Disorder (rrMDD). This study involved 50 rrMDD patients and 25 healthy control participants, and it aimed to understand neural regulation problems in daily life. The findings indicated that rrMDD patients exhibited lower activation and different temporal dynamics in certain brain regions during passive attendance of emotional material compared to health controls. Additionally, dysfunctional regulation strategies in daily life and lower effects were observed in rrMDD patients. CBT has been shown to impact the neural correlates of ER in individuals with MDD. Prior research has demonstrated that CBT can lead to changes in brain activity, particularly in areas associated with ER, such as the prefrontal cortex and limbic system. Specifically, CBT has been linked to increased prefrontal cortex activity and decreased limbic system activity, indicating improved regulation of emotions and mood. These changes in neural activity may contribute to the effectiveness of CBT in treating depression and preventing relapse (van Kleef et al. 2022, 1-2). In summary, this study examined the neural correlates of ER in individuals with Major Depressive Disorder and has provided valuable insights into the enduring irregularities in emotional regulation observed even after remission, encompassing both the management of positive and negative emotions. Furthermore, the impact of CBT on the neural correlates of ER highlights the potential for therapeutic interventions to modulate brain activity and improve ER in individuals with MDD.

A third study, by Schweizer and colleagues (2020), discussed the relationship between effective control, ER, and mental health, particularly in adolescent development. It highlights the importance of understanding the typical development of affective control and its association with ER tendencies, as poor affective control is linked to mental health problems in adolescence. The

review emphasizes the need for longitudinal research to explore the co-development of affective control as a cognitive building block of successful ER, with good affective control associated with fewer ruminative tendencies and more successful reappraisal in older adolescents. Lower use of habitual suppression is also linked to updating capacity. The document also discusses the potential mental health benefits associated with improved effective control (Schweizer, Gotlib, and Blakemore 2020, 80-84). The review acknowledges that longitudinal research is currently lacking and that the relative contribution of different facets of affective control compared to cool cognitive control is poorly understood. It also emphasizes the need for psychometric investigations to improve mechanistic understanding of their relative contributions. The discussion emphasized that CBT can lead to changes in neural activity and connectivity in brain regions involved in ER. For example, it demonstrated that CBT can lead to increased prefrontal cortex activity and reduced amygdala reactivity in response to emotional stimuli. Additionally, CBT has been associated with changes in connectivity between prefrontal brain regions and emotion-processing regions, which may contribute to improved ER.

Concerns and Limitations

While the benefits conferred by neuroimaging techniques in studying neural correlates of CBT interventions for MDD significantly advance the understanding of neural mechanisms, it is imperative to acknowledge the limitations associated with these studies. Notably, several studies primarily rely on self-report instruments for treatment outcome assessment, potentially introducing bias and subjective interpretations into the analysis. Moreover, comparisons between CBT-ERT (Enhanced Relapse Prevention) and control conditions sometimes involve degraded conditions such as sports and occupational therapy, which may confound results and obscure the specific effects of CBT on neural correlates. These limitations underscore the importance of adopting rigorous methodologies to minimize confounding factors and enhance the reliability of findings. One approach commonly employed to address these concerns is the utilization of Randomized Controlled Trials (RCTs), which allow researchers to control for variables that may influence treatment outcomes, including participant characteristics and environmental factors. By randomly assigning participants to treatment and control groups, RCTs reduce the risk of bias and increase the internal validity of the study, thereby providing more robust evidence regarding the neural mechanisms underlying CBT interventions for MDD. Despite the advancements proposed by these methodological innovations, ongoing efforts are needed to address remaining challenges and further elucidate the complex interplay between neural processes and therapeutic interventions in MDD.

A second limitation is that neuroimaging studies often recruit participants with specific demographic or clinical characteristics, such as age, which may not fully represent the diverse population of individuals with MDD. For example, studies may have a majority of participants as adolescents, potentially limiting the applicability of findings to the elderly who may experience depression differently or respond differently to CBT. The inclusion criteria used in neuroimaging studies examining the effects of CBT on neural correlates may exclude certain subgroups of

individuals with MDD, such as those with specific treatment preferences of cultural backgrounds. This selective sampling approach limits the generalizability of findings and may overlook important factors that influence treatment response in real-world clinical settings. The duration and intensity of CBT interventions administered in research settings may also differ from those delivered in routine clinical practice, further complicating the extrapolation of neuroimaging findings to everyday treatment contexts. Moreover, variations in illness severity among participants can significantly impact neural responses to CBT and contribute to treatment outcomes. Individuals with more severe depression may exhibit distinct neural signatures compared to those with milder forms of the disorder, potentially influencing how they respond to CBT interventions. For example, individuals with severe depression may demonstrate hypoactivation in prefrontal regions associated with cognitive control and emotion regulation, which could affect their ability to engage with CBT techniques effectively. Conversely, individuals with milder depression may show patterns of neural activity in response to CBT, reflecting variations in symptom presentation and underlying neural circuitry.

Despite the advancement proposed by existing methodologies, ongoing efforts are needed to address remaining challenges and further elucidate the complex interplay between neural processes and therapeutic interventions in MDD. For example, one notable concern is the variability in treatment response observed among individuals undergoing CBT. While some may exhibit substantial improvement in depressive symptoms, others may show limited or no response. The variability in treatment response observed among individuals undergoing CBT for MDD underscores the influence of individual differences, including genetic predispositions, neurobiological profiles, and psychosocial factors (Gautam et al. 2020, S224). Genetic factors influence an individual's responsiveness to CBT, as genetic variations may affect neurotransmitter function or stress response systems, thus impacting treatment outcomes (Lester and Eley 2013, 1). Additionally, variations in neurobiological profiles, such as differences in brain structure and function, may contribute to differential responses to CBT (Gautam et al. 2020, S224-225). For instance, individuals with alterations in brain regions implicated in emotion regulation, such as the prefrontal cortex or amygdala, may exhibit distinct responses to CBT interventions (Alipour, Arefnasab, and Babamahmoodi 2011, 114-115). Furthermore, psychosocial factors, including past trauma, social support networks, and personality traits, can shape an individual's engagement with therapy and influence treatment effectiveness (Alipour, Arefnasab, and Babamahmoodi 2011, 114-115).

Critics argue that the efficacy of CBT in treating MDD may be overstated, as other therapeutic modalities, like interpersonal therapy (IPT), may offer comparable or even superior outcomes for certain individuals. IPT focuses on addressing interpersonal conflicts and improving social functioning, targeting psychosocial factors that contribute to depressive symptoms. For individuals whose depression is influenced by relational dynamics or interpersonal stressors, IPT offers distinct advantages over CBT by directly addressing these underlying issues (Rajhans et al. 2020, S202). While CBT has consistently demonstrated efficacy in reducing depressive symptoms and preventing relapse, its relative effectiveness compared to

alternative treatments may vary depending on individual characteristics and preferences. Although some individuals may prefer the structured and skill-based approach of CBT, finding it empowering to actively engage in identifying and modifying their thought patterns and behaviours, others may resonate more with the interpersonal focus of alternative therapies, leading to differential treatment preferences and outcomes. This highlights the importance of considering individual factors when developing treatment plans

Future Directions

In the field of CBT for MDD and ER, several avenues remain unexplored, offering various opportunities for future research endeavours. One critical aspect that needs further investigation pertains to the precise mechanisms underlying the efficacy of CBT in treating MDD and enhancing ER skills. While existing studies have provided valuable insights into the cognitive restructuring techniques and emotion regulation strategies employed in CBT, there remains ambiguity regarding the specific neurobiological and psychological processes affecting treatment outcomes. Unravelling these intricate mechanisms could offer profound implications for refining therapeutic approaches, potentially leading to more targeted and effective interventions for individuals grappling with MDD. For instance, understanding how neural circuits implicated in the emotional processing and regulation are modulated by CBT interventions, or how individual differences in cognitive biases and emotional reactivity influence treatment response, could contribute to the underlying mechanisms driving symptom remission. Exploring the role of therapeutic factors such as the therapeutic alliance, treatment adherence, and cognitive flexibility in CBT outcomes may provide valuable insights into optimizing treatment delivery and tailoring interventions to individual patient needs. By delving deeper into these unanswered questions, researchers can not only enhance our understanding of the mechanisms underlying CBT efficacy but also pave the way for more personalized and effective interventions for individuals with MDD.

Advancing our understanding of the long-term effects of CBT on MDD symptomatology and ER capacities is imperative. While short-term outcomes of CBT for MDD have been well-documented, there remain areas that have not yet been investigated such as the durability of treatment effects over extended periods. Longitudinal investigations can be beneficial to the durability and sustainability of treatment that is gained over time: spanning months or years. Furthermore, such studies could support whether the skills acquired during CBT, such as cognitive restructuring and ER, are maintained over time and contribute to sustained improvements in mental health outcomes. Additionally, exploring how individual differences in treatment response evolve could inform the development of tailored interventions that address the evolving needs of patients throughout their recovery journey. Investigating the potential impact of contextual factors, such as life stressors or changes in social support networks, on the long-term efficacy of CBT could enhance our understanding of the real-world effectiveness of this therapeutic approach. By addressing these knowledge gaps, researchers can provide

clinicians with evidence-based guidelines for the optimal timing and duration of CBT interventions, ultimately improving the long-term prognosis for individuals living with MDD.

In addressing the multifaceted nature of MDD, a personalized medicine approach is essential in future research. This tailored intervention to specific patient characteristics is crucial for optimizing outcomes. Through the utilization of advanced analytical techniques such as machine learning algorithms and predictive modeling, research can identify patient-specific biomarkers and psychosocial factors predictive of treatment response. By integrating this personalized approach into clinical practice, healthcare professionals can deliver more targeted and effective interventions, minimizing trial-and-error approaches and maximizing treatment efficacy. Subsequently, this personalized medicine paradigm has the potential to revolutionize the field of mental health care, offering a pathway toward precision medicine in psychiatry. By aligning treatment strategies with individual patient needs, clinicians can foster greater patient engagement, enhance treatment adherence, and ultimately alleviate the burden of MDD on both patients and healthcare systems. Thus, embracing these future directions in research not only expands the understanding of MDD pathology and treatment mechanisms but also holds the promise of transforming clinical care, offering hope for individuals navigating the complexities of depressive disorders.

Moving forward, recommendations for future research directions to enhance the public understanding of the neurobiological basis of MDD and improve treatment strategies include incorporating advanced neuroimaging techniques such as fMRI and EEG to elucidate the temporal dynamics of neural changes associated with CBT. Moreover, longitudinal studies with larger sample sizes are needed to investigate the durability of neural alterations induced by CBT and their relationship to long-term clinical outcomes. Furthermore, integrating multimodal approaches combining neuroimaging with other biomarkers, such as genetic, hormonal, and inflammatory markers, may provide a more comprehensive understanding of underlying mechanisms driving treatment response in MDD.

Conclusion

The analysis of CBT's impact on neural correlates in MDD has revealed several key findings and insights. Firstly, neuroimaging studies have consistently demonstrated alterations in brain regions implicated in ER, such as the prefrontal cortex, amygdala, and hippocampus, among individuals with MDD (Yuan et al. 2022, 6). Importantly, CBT interventions have been associated with normalization of activity in these regions, suggesting a neurobiological mechanism underlying its therapeutic effects. Additionally, longitudinal studies have indicated changes in functional connectivity patterns within neural circuits involved in emotion processing following CBT treatment, further supporting the notion of neuroplasticity and neural adaptation associated with psychological interventions in MDD (Yuan et al. 2022, 6).

Individual differences in neural responses to CBT underscore the necessity of personalized treatment strategies tailored to patients' specific neurobiological profiles. The multifaceted nature of depression suggests that a singular universal treatment approach may not

be effective for all individuals. Instead, having personalized treatment plans considering individuals' needs, preferences, and factors contributing to depression is likely to yield optimal outcomes. Neuroimaging studies examining the effects of CBT on neural correlates in MDD provide promising insights into therapeutic mechanisms, emphasizing the integration of neurobiological perspectives into clinical practice.

Works Cited

- Bains, N., and S. Abdijadid. "Major Depressive Disorder." Last updated April 10, 2023. In *StatPearls*. Treasure Island, FL: StatPearls Publishing, 2024. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK559078/>.
- Villarroel, Maria A., and Emily P. Terlizzi. "Products - Data Briefs - Number 379 - September 2020." Centers for Disease Control and Prevention, 2020. <https://www.cdc.gov/nchs/products/databriefs/db379.htm>.
- World Health Organization (WHO). "Depressive Disorder (Depression)." 2023. <https://www.who.int/news-room/fact-sheets/detail/depression>.
- Schweizer, Susanne, Ian H. Gotlib, and Sarah-Jayne Blakemore. "The Role of Affective Control in Emotion Regulation during Adolescence." *PubMed*, 2020. <https://pubmed.ncbi.nlm.nih.gov/31961183/>.
- Gautam, Manaswi, Adarsh Tripathi, Deepanjali Deshmukh, and Manisha Gaur. "Cognitive Behavioral Therapy for Depression." *Indian Journal of Psychiatry* 62, no. Suppl 2 (2020): S223-S229. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_772_19.
- Xie, Musi, Ying Huang, Wendan Cai, and Jiang Y. "A Theory of the Neural Mechanisms Underlying Negative Cognitive Bias in Major Depression." *Frontiers in Psychiatry* 15 (2024): 1348474. <https://doi.org/10.3389/fpsyt.2024.1348474>.
- Xie, Musi, Ying Huang, Wendan Cai, and Jiang Y. "A Theory of the Neural Mechanisms Underlying Negative Cognitive Bias in Major Depression." *Frontiers in Psychiatry* 15 (2024): 1348474. <https://doi.org/10.3389/fpsyt.2024.1348474>.
- Phillips, M. L. "Understanding the Neurobiology of Emotion Perception: Implications for Psychiatry." *Cambridge University Press*, 2003. <https://doi.org/10.1192/bjp.182.3.190>.
- Tartt, Alexandria N., Madeline B. Mariani, Rene Hen, J. J. Mann, and Maura Boldrini. "Dysregulation of Adult Hippocampal Neuroplasticity in Major Depression: Pathogenesis and Therapeutic Implications." *Molecular Psychiatry* 27, no. 6 (2022): 2689-2699. Accessed April 20, 2024. <https://doi.org/10.1038/s41380-022-01520-y>.
- Rădulescu, I., A. M. Drăgoi, S. C. Trifu, and M. B. Cristea. "Neuroplasticity and Depression: Rewiring the Brain's Networks through Pharmacological Therapy (Review)." *Experimental and Therapeutic Medicine* 22, no. 4 (2021): 1131. <https://doi.org/10.3892/etm.2021.10565>.
- Cui, L., S. Li, S. Wang, et al. "Major Depressive Disorder: Hypothesis, Mechanism, Prevention and Treatment." *Signal Transduction and Targeted Therapy* 9 (2024): 30. <https://doi.org/10.1038/s41392-024-01738-y>.
- Apps, M. A., M. F. Rushworth, and S. W. Chang. "The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others." *Neuron* 90, no. 4 (2016): 692-707. <https://doi.org/10.1016/j.neuron.2016.04.018>.
- Pizzagalli, Diego A., and Angela C. Roberts. "Prefrontal Cortex and Depression." *Neuropsychopharmacology* 47, no. 1 (2022): 225-246. Accessed April 21, 2024. <https://doi.org/10.1038/s41386-021-01101-7>.

- Singh, Kuljeet A., and Vikas Dhikav. "Hippocampus in Health and Disease: An Overview." *National Library of Medicine*, 2012.
https://journals.lww.com/annalsofian/fulltext/2012/15040/hippocampus_in_health_and_disease__an_overview.2.aspx.
- Nakao, Mutsuhiro, Kentaro Shiotsuki, and Nagisa Sugaya. "Cognitive–Behavioral Therapy for Management of Mental Health and Stress-related Disorders: Recent Advances in Techniques and Technologies." *BioPsychoSocial Medicine* 15 (2021). Accessed April 21, 2024. <https://doi.org/10.1186/s13030-021-00219-w>.
- Yen, Chiahui, Chia-Li Lin, and Ming-Chang Chiang. "Exploring the Frontiers of Neuroimaging: A Review of Recent Advances in Understanding Brain Functioning and Disorders." *Life* 13, no. 7 (2023): 1472. <https://doi.org/10.3390/life13071472>.
- Kotoula, Vasileia, Jennifer W. Evans, and Claire E. Punturieri. "Review: The Use of Functional Magnetic Resonance Imaging (fMRI) in Clinical Trials and Experimental Research Studies for Depression." *Frontiers in Neuroimaging* 2 (2023). Accessed April 21, 2024. <https://doi.org/10.3389/fnimg.2023.1110258>.
- Rangaswamy, Madhavi, and Bernice Porjesz. "From Event-Related Potential to Oscillations: Genetic Diathesis in Brain (Dys)Function and Alcohol Dependence." *Alcohol Research & Health* 31, no. 3 (2008): 238-242. Accessed April 21, 2024. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3860487/>.
- Liu, X., H. Zhang, Y. Cui, T. Zhao, B. Wang, X. Xie, S. Liang, S. Sha, Y. Yan, X. Zhao, and L. Zhang. "EEG-based Major Depressive Disorder Recognition by Neural Oscillation and Asymmetry." *Frontiers in Neuroscience* 18 (2024): 1362111. <https://doi.org/10.3389/fnins.2024.1362111>.
- Finnema, Sjoerd J., Mika Scheinin, Mohammed Shahid, Jussi Lehto, Edilio Borroni, Benny Bang-Andersen, Jukka Sallinen, et al. "Application of Cross-species PET Imaging to Assess Neurotransmitter Release in Brain." *Psychopharmacology* 232 (2015): 4129-4157. Accessed April 21, 2024. <https://doi.org/10.1007/s00213-015-3938-6>.
- Poli, A., A. Pozza, G. Orrù, C. Conversano, R. Ciacchini, D. Pugi, N. L. Angelo, L. L. Angeletti, M. Miccoli, and A. Gemignani. "Neurobiological Outcomes of Cognitive Behavioral Therapy for Obsessive-Compulsive Disorder: A Systematic Review." *Frontiers in Psychiatry* 13 (2022): 1063116. <https://doi.org/10.3389/fpsyt.2022.1063116>.
- Berking, Matthias, David Ebert, Pim Cuijpers, and Stefan G. Hofmann. "Emotion Regulation Skills Training Enhances the Efficacy of Inpatient Cognitive Behavioral Therapy for Major Depressive Disorder: A Randomized Controlled Trial." *Karger*, 2013. <https://doi.org/10.1159/000348448>.
- van Kleef, R. S., J. C. Marsman, E. van Valen, C. L. H. Bockting, A. Aleman, and M. J. van Tol. "Neural Basis of Positive and Negative Emotion Regulation in Remitted Depression." *NeuroImage: Clinical* 34 (2022): 102988. <https://doi.org/10.1016/j.nicl.2022.102988>.
- Lester, Kathryn J., and Thalia C. Eley. "Therapygenetics: Using Genetic Markers to Predict Response to Psychological Treatment for Mood and Anxiety Disorders." *Biology of

- Mood & Anxiety Disorders* 3, no. 1 (2013): 4. <https://doi.org/10.1186/2045-5380-3-4>. Published February 7, 2013.
- Alipour, Abbas, Zahra Arefnasab, and Ahmad Babamahmoodi. "Emotional Intelligence and Prefrontal Cortex: A Comparative Study Based on Wisconsin Card Sorting Test (WCST)." **Iranian Journal of Psychiatry and Behavioral Sciences** 5, no. 2 (2011): 114-119.
- Rajhans, Pallavi, Gagan Hans, Vinay Kumar, and Rakesh Kumar Chadda. "Interpersonal Psychotherapy for Patients with Mental Disorders." **Indian Journal of Psychiatry** 62, Suppl 2 (2020): S201-S212. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_771_19.
- Yuan, Shiting, Huiqin Wu, Yun Wu, Huazhen Xu, Jianping Yu, Yuan Zhong, Ning Zhang, Jinyang Li, Qianwen Xu, and Chun Wang. "Neural Effects of Cognitive Behavioral Therapy in Psychiatric Disorders: A Systematic Review and Activation Likelihood Estimation Meta-Analysis." **Frontiers in Psychology** 13 (2022). Accessed April 21, 2024. <https://doi.org/10.3389/fpsyg.2022.853804>.

Innovations for *In Vivo* Gene Therapy Delivery and the AI-Driven Future of Genetic Medicine By Norman Lwin

Abstract

Throughout the last century, gene therapy has largely relied on *ex vivo* delivery using viral vectors like lentivirus and adenovirus. However, with the development of innovative delivery modalities such as lipid nanoparticles, polymeric nanoparticles, and virus-like particles, the transition from laborious *ex vivo* delivery to cost-effective *in vivo* delivery has become widely spread. These various nanocapsules have been carefully engineered to maximize efficiency, precision, and safety to treat a multitude of genetic conditions. Nanocapsules, intricately integrated with CRISPR-Cas9 systems, are propelling gene editing into an unprecedented era. Cutting-edge advancements in delivery technologies, combined with innovative gene editing techniques such as base editing and prime editing, hold transformative potential for tackling previously incurable diseases like Cystic Fibrosis. Meanwhile, the integration of AI is revolutionizing the field, enabling precise protein structure prediction and the design of optimal nanocarriers. These breakthroughs are not only enhancing the effectiveness of established therapies but also paving the way for groundbreaking treatments, pushing the boundaries of what's possible in gene therapy. While gene editing shows much promise, limitations still arise concerning safety and efficacy, pushing efforts to refine the established therapies. In order to materialize the successes seen in *ex vivo* gene therapy, the field needs to develop safe and effective therapies to deliver novel DNA-modulating modalities to specific cell types of interest where the disease manifests. In this review, we will discuss the novel delivery systems to move gene therapies toward *in vivo* applications, the current limitations of gene therapy, and the novel payloads required to treat these diseases.

Introduction

Genetic disorders and diseases are among the most complicated and fatal diseases in modern-day medicine particularly due to the intricacy of the human genome. The human genome is all the genetic information stored in the body, consisting of DNA strands coiled and packed tightly into chromosomes. Each segment of the DNA is known as a gene which is split into protein-coding exons, and non-protein-coding introns. It is estimated that there are 20,000 exons in the human genome which comprises no more than 3% of the genome as a whole (**Huntington F. Willard 2017**). Thus, the intricacy of the genetic system leads to extremely complex challenges when error occurs. Mutations typically occur during DNA replication before cell division (somatic mutations) or can be inherited if they occur in sperm or egg cells (germ-line mutations). Cells will attempt to use their DNA-repairing enzymes to correct the mutation, however many times, this can lead to further error, potentially deleting or inserting new segments. Thus, with the unpredictability of genetic disorders, gene therapies strive to manipulate the human genome to overcome their fickle nature precisely.

Genetic diseases, often inherited and beyond patients' control, drive the urgent need for

advancements in gene therapy to alleviate suffering. With the WHO estimating that 1 in 100 people—70-80 million globally—are affected by monogenic diseases, the demand for effective solutions has never been clearer. Traditional treatments for genetic disorders came from allogenic donations, such as harvesting healthy blood cells to donate to patients. However, this raised many concerns. Graft vs Host Disease, a risk of infection, and overall, the inconvenient process of having to find a healthy donor made it an unreliable treatment. Thus, the development of early key genetic therapies offered a solution. TALENS and Zinc Fingers, which were nuclease-based editing tools, were primarily used in *ex-vivo* conditions to edit hematopoietic stem cells **(Thomas Gaj 2013)**. The tools demonstrated high success yet the *ex vivo* processes required inefficient processes that needed improvement. The field of gene therapy is embracing advanced *in-vivo* delivery systems to deploy cutting-edge tools like CRISPR-Cas9, base editing, and prime editing, driving progress against diseases like Cystic Fibrosis and Huntington's..

Therapeutics and Delivery Challenges

Drug delivery systems (DDS) are methods for administering a pharmaceutical compound to achieve a therapeutic effect. For decades, most medicine was delivered systemically resulting in unpredictability and inefficiency. Conventional small-molecule drugs are released immediately which requires frequent administration and hinders their ability to reach target cells **(H. Park 2021)**. Additionally, concerns regarding drug solubility and stability rose as about 90% of preclinical drugs have low solubility impeding their ability to circulate throughout the body **(A. Vargason 2021)**. Therefore DDS have been engineered to optimize the delivery of therapeutic agents to specific sites, attempting to eliminate off-target effects and inefficacies. Peptides and proteins allow for more precise targeting because they can bind more specifically to receptors compared to small molecules. However, due to their large and complex structures, they are prone to protease degradation, environmental conditions like pH and temperature, and immune responses. Other well-known therapeutics include antibodies that can bind to specific antigens and neutralize them by inducing phagocytosis, and nucleic acids which can precisely control gene regulation and expression. Despite their useful functions, antibiotics risk inducing adverse effects such as autoimmune diseases due to their foreign nature, and nucleases quickly degrade nucleic acids before achieving their intended function **(A. Vargason 2021)**.

Evidently, these therapeutic agents were missing a key factor in optimizing their treatment and capabilities. Thus, with the discoveries of the liposome, drug delivery systems would experience a translational revolution resulting in precise, innovative, and highly functional delivery vehicles. In 1961, Alec Bangham and colleagues were studying the properties of phospholipids and their ability to form lipid bilayers which had not been fully understood. They discovered that when phospholipids were dispersed in water, they formed spherical structures with a lipid bilayer. The hydrophobic tails of the phospholipid faced outward, away from the water, while the hydrophilic heads face inwards, towards the water. Using electron microscopy, they were able to confirm that the structures they found were vesicles with a lipid bilayer and noted that the structure had properties characteristic of cell membranes: permeability and a

bilayer structure. This would facilitate the study of cell membranes and their interactions with other molecules, a fundamental biological concept in all living organisms (**A.D Bangham 1965**). Later on, Dr. Greg Gregoriadis discovered that the liposome would be able to serve a purpose much more useful than as a study tool. In 1971, Gregoriadis hypothesized that the spherical structure could potentially encapsulate drugs or enzymes to deliver them to specific parts of the body. Thus, proteins were mixed in a phospholipid and aqueous solution to evaluate the efficacy of encapsulation. The team found that the liposome-encapsulated a high protein concentration and successfully protected the cargo from proteolytic enzymes (**G. Gregoriadis 1972**).

Following Gregoriadis' studies, Dr. Yechezkel Barenholz and Dr. Alberto Gabizon initiated the first clinically approved liposomal treatment known as Doxil in 1995. While doxorubicin had been successful as a chemotherapy agent, the researchers wanted to use the liposomes' novel applications to reduce adverse effects such as cardiotoxicity and off-targeting. The liposomal structure was intricately engineered for optimization. The doxorubicin was encapsulated within the liposome's aqueous core and coated with PEG. Due to the incredibly small size of the liposome and the high dosage of doxorubicin required, a transmembrane gradient of ammonium sulfate was required to maintain stability. The ammonium ions within the liposome protonate the doxorubicin, allowing for an electrostatically attracted package. Once the liposomes have localized, the acidic environment of the tumor environment encourages the liposome shell to break down and release its cargo, producing a controlled release of the doxorubicin to be taken into the cell via endocytosis and inducing cell apoptosis. Doxil exploits both passive and active targeting systems to optimize its treatment efficacy. Thus, once the design had passed preclinical studies in animals, the treatment had reached its third phase where it was used for patients with ovarian cancer and AIDS-related Kaposi's sarcoma (ARKS) (**Y. Barenholz 2012**). The study presented significant results with high antitumor activity in ARKS and ovarian cancer. All treatments demonstrated effective delivery and no toxicities associated (**A. Gabizon 2001**). With its successes and safety, Doxil became the first liposomal delivery system approved by the FDA in 1995, bringing even more promise and effort into developing drug delivery systems.

Delivery Vehicles Currently in Clinical Use

Delving into the field of gene therapy, viral vectors have dominated due to their ability to infect cells. By carrying a nucleic acid, viral vectors can easily integrate into the cell's genome, facilitating the study of gene function, and gene therapy. Adenoviruses, an example of one, were first discovered in the 1950s when researching respiratory illnesses. A key characteristic of this family of viruses was that they were able to infect both dividing and nondividing cells, encouraging a wide scope of infection. As molecular biology advanced, researchers realized that these viruses could be genetically modified to carry foreign genes and be rid of their replication-responsible genes. They were also found to be highly immunogenic, stimulating strong immune responses and highlighting their potential as a vaccine. Thus, with its unique properties, adenoviruses quickly converted into a strong research tool. (**Doerfler, 2010**)

In contrast to adenoviruses, researchers discovered a viral vector that could integrate into the genome because it carries RNA rather than DNA. Lentiviruses, most commonly originating from HIV, could infect both dividing and non-dividing cells and integrate into the genome smoothly and stably which was not seen in retroviruses. Thus, with these properties, researchers focused on applying lentiviral vectors for gene therapies, as they were more stable and had a wider range of applications than retroviruses which were limited to infecting only dividing cells. Lentivirus vectors consist of three main genes: a gag, which encodes structural proteins; pol, which encodes the reverse transcriptase enzyme, and the env which encodes the viral envelope glycoprotein, facilitating the virus' integration into the genome. Previously, retroviruses had been used to study gene function and gene therapy due to their natural ability to reverse-transcribe their RNA to DNA and infect other genomes. It was a sufficient research tool however, had many limitations including its infection range, and potential toxicities. Thus, lentiviral vectors were used in the same way as they offered a solution to the retroviruses' limitations. Novel applications of lentivirus include cell engineering like CAR-T for cancer immunotherapy, and gene delivery to restore functionality in faulty cells, and can even be used to create transgenic animal models. Overall, lentiviral vectors have proven to be a successful gene delivery system (M. Milone 2021).

While both adenoviruses and lentiviruses have been used as viral vectors for gene therapy, they mainly differentiate in their structure and stability. Adenoviruses are non-enveloped viruses with a double-stranded DNA, while lentiviruses are enveloped viruses with a single-stranded RNA. Because of this, adenoviruses don't integrate into the host's genome, rather, release their DNA into the nucleus once it has been taken in through endocytosis. Lentiviruses integrate into the host's genome and reverse transcribe its RNA into DNA which integrates into the nucleus. The difference in these methods causes adenoviruses to have transient expression, meaning it is temporary, while lentiviruses have stable long-term expression. Both have also been crucial for studying genes in general however adenoviruses have a higher focus on vaccines as they are highly immunogenic and lentiviruses focus more on cell engineering as they induce long-term gene expression which is essential for immunotherapy. Overall, both have been crucial in advancing the field of gene therapy, allowing for the technologies and advancements we have made today.

The application of lentiviral vectors (LVs) in clinical trials for *in vivo* gene therapies has brought up issues related to genotoxicity, immune responses, and the expensive nature of manufacturing (Milone and O'Doherty, 2018). With many concerns, any of these challenges could potentially restrict the use of LVs in *in vivo* gene editing approaches. Similar to adenoviruses, AAVs also contain a single strand of DNA. However, they are much smaller in size typically ranging from 20-25 nms while AVs range from 70-90 nms. As a result, AVs can carry larger payloads, however, are less precise and efficient for gene therapy. AAVs are less immunogenic than AVs, stimulating less immune responses; can infect dividing and non-dividing cells similar to LVs, broadening their tropism; and can integrate into the host genome unlike the AVs. Due to its ability to integrate into the genome, AAVs are used for long-term expression

rather than the transient expression offered by the AVs. The low immunogenicity allows AAVs to reach their targets more successfully while the high immunogenicity allows AVs to better serve as vaccines to prompt immune responses (**M. F. Naso**). The three main categories of viral vectors are comprehensively compared in Figure 1, highlighting the weaknesses and strengths of each one.

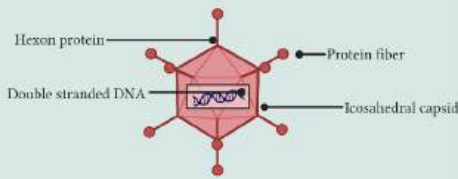
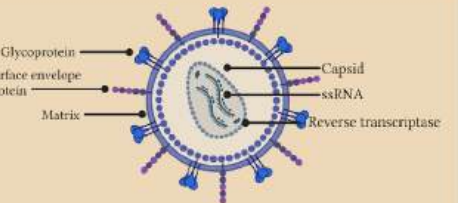
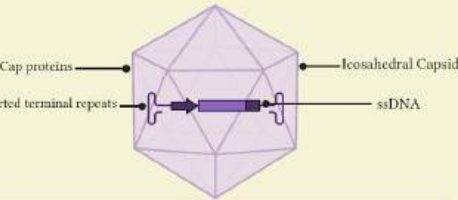
	Structure	Overview
Adenovirus Cargo Size: 8-10 kb	 Diagram illustrating the structure of an Adenovirus. It shows a red icosahedral capsid with protein fibers extending from its vertices. Inside the capsid, double-stranded DNA is visible. Labels include: Hexon protein, Protein fiber, Double stranded DNA, and Icosahedral capsid.	Adenoviruses are <i>non-enveloped</i>, <i>double stranded DNA</i> viruses. Its <i>icosahedral capsid</i> made up of <i>hexon and fiber proteins</i> facilitate their entry into cell. Upon cell entry, the DNA is <i>transcribed into mRNA and replicated</i>, taking over the protein producing mechanism of the host cell. They are highly immunogenic and can infect cells easily making them <i>valuable vaccine and research tools</i>.
Lentivirus Cargo Size: 8-10 kb	 Diagram illustrating the structure of a Lentivirus. It shows a blue spherical virus with a surface envelope protein and glycoprotein spikes. Inside, there is a matrix, a capsid, and ssRNA. Labels include: Glycoprotein, Surface envelope protein, Matrix, Capsid, ssRNA, and Reverse transcriptase.	Lentiviruses contain a <i>ssRNA</i> in the <i>capsid</i> which is <i>reverse transcribed</i> upon cell entry and integrated into the host genome. Three main genes are: <i>gag gene</i> encodes structural proteins, <i>pol gene</i> encodes enzymes for viral replication, and <i>env gene</i> encodes glycoproteins for cell entry Lentiviruses can infect both <i>dividing and non-dividing</i> cells, safely without producing any immunogenic proteins. Very strong for <i>gene knockout and editing</i>.
AAV (Adeno-associated virus) Cargo Size: 4.7 kb	 Diagram illustrating the structure of an AAV (Adeno-associated virus). It shows a purple icosahedral capsid with cap proteins and inverted terminal repeats. Inside, ssDNA is visible. Labels include: Cap proteins, Inverted terminal repeats, Icosahedral Capsid, and ssDNA.	AAVs are <i>non-enveloped</i> viruses with a <i>icosahedral capsid</i> and a <i>ssDNA</i>. AAV genome includes three main genes: <i>rep gene</i> encodes enzymes essential for viral replication, <i>cap gene</i> encodes structural proteins for the viral capsid, and <i>aap gene</i> which encodes proteins for assemble of AAVs AAV transports ssDNA to the nucleus where it is transcribed and replicated to infect both dividing and non-dividing cells. AAV is <i>less immunogenic</i> than adenoviruses making it better suited for long term therapies.

Figure 1: Mechanism and Structure of Viral Delivery Modalities

Novel Delivery Vehicles in Development for Gene Therapy

Over the past decade, LNPs have evolved from liposomes and have demonstrated impressive pharmacological performance and therapeutic promise. While closely related, LNPs are essentially a newer generation of liposomes mainly differentiating in their stability and structure. LNPs are typically easier to produce and contain a solid core, where the lipids are melted to form a liquid which later cools containing the therapeutic drug to solidify and encapsulate the drug, while liposomes form an aqueous core due to their dependency on its lipid bilayer. The solid core reduces the chances of drug leakage and degradation enhancing its stability and can solubilize lipophilic drugs, improving their ability to dissolve and absorb into the bloodstream. Furthermore, the solid core can be engineered to control drug release improving its targeting capabilities. Overall, modern LNPs are more precise, effective, and cost-efficient

forms of liposomes. Despite this, liposomes are still widely used due to their continued proof of reliability.

Contrasting the liposome's reliance on natural interactions, polymeric nanoparticles are designed and engineered. Constructed by nano-sized biodegradable and biocompatible polymers, these polymers can be both natural (e.g., chitosan, gelatin) and synthetic (e.g., polylactic acid (PLA)) allowing for a versatile means of production. Two main strategies known as solvent evaporation and emulsification are used to produce polymeric nanoparticles. Polymers can be dissolved in an organic solvent, alongside a surfactant to facilitate mixing, and are mixed using high-speed homogenization or ultrasonication. This creates tiny droplets in the mixture that form into solid nanoparticles encapsulating the drug once the solvent has evaporated. Emulsification involves creating an oil-in-water emulsion, combining a solvent mixed with the polymer and drug. When a large amount of water is added, the insoluble solvent begins to diffuse out, forming drug-encapsulated nanoparticles (Zielinska et al., 2020). Thus, through this process forms the permeable, precise, biodegradable, and efficient DDS that the polymeric nanoparticle can be used as (H. Bhardwaj and Jangde, 2023).

Among cutting-edge drug delivery systems, viral-like particles (VLPs) are one of the leading forces in this field. These particles are engineered to mimic viral particles without containing any genetic material through self-assembly from proteins encoding viral structures like capsids, cores, or envelopes. The first step involves a process called gene cloning where the genes encoding specific viral structures are selected and amplified through Polymerase Chain Reaction, creating multiple copies of the selected gene. An expression vector is chosen next, typically plasmid DNA and viral vectors, to deliver the gene into the host cell, replacing a strand of DNA and being strung together by ligase. DNA sequences that promote transcription are also incorporated into the complex. After entering a cell's genome, the DNA goes through transcription and translation producing the viral protein structures. VLPs can then be functionalized with ligands or polymers and encapsulate cargo (Mejía-Méndez et al., 2022). Due to their similarities to viruses, VLPs have high immunogenicity making them useful for vaccines and stimulating immune responses. VLPs have been enhanced to carry prime editing complexes. Previously, they have had low success in editing genes and were inefficient. However, after making advancements in the sgRNA, mismatch repair evasion systems, and the elimination of the Gag polyprotein which allows for higher structural stability in the VLPs, they experienced a 170-fold increase in editing efficiency and were able to correct 4 base pairs in the retina of mice, restoring partial visual function and achieving editing rates comparable to adenoviruses. These engineered VLPs show promise as cutting-edge delivery vehicles (M. An et al., 2024). Recent efforts can be seen through the new VLP delivery platform developed by Nvelop Therapeutics known as DLVR-M. The DLVR-M is a novel platform for protein delivery composed of an envelope-pseudotyped vesicle that is fused to a PH domain, a protein domain that facilitates the targeting of certain lipids. Furthermore, this delivery platform can efficiently encapsulate gene editing technologies such as CRISPR-Cas9, ABEs, and PEs. In one specific trial, the DLVR-M particle was able to cause a 75% reduction in the PCSK9 gene, responsible for producing LDL

cholesterol, a type of cholesterol that can build plaques in circulation. Clearly, the new VLP platform shows much promise in the field of gene therapy. The structure and mechanism for each developing delivery modality are summarized and compared in Figure 2.

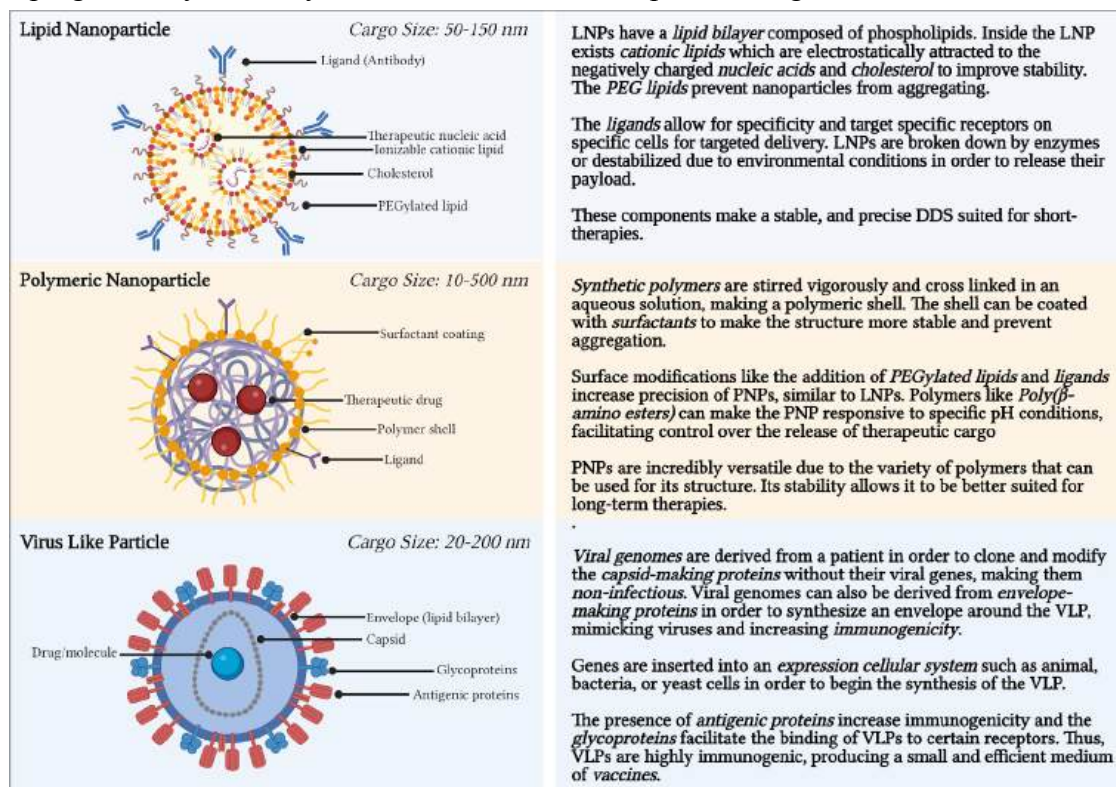


Figure 2: Novel non-viral delivery modalities that are widely used

Gene Editors in Development

With the establishment and ongoing research of advanced and efficient delivery modalities, all areas of medicine are revolutionized. Gene editing therapies and technologies specifically will find these modalities appealing, transitioning from a demanding and suboptimal *ex vivo* process to a cost-effective and efficient in-vivo method that does not risk mutations. Gene editing has revolutionized the field of gene therapy, a once primitive tool using Zinc Fingers or TALENS to edit cells *ex vivo* in a laborious, slow, and inefficient process that could never be translated to *in vivo* gene therapy. However, since the discovery of Clustered regularly interspaced short palindromic repeat (CRISPR) - Cas systems, the need for these primitive tools has diminished.

In 1987, Yoshizumi Ishino and colleagues were dissecting phosphate metabolism genes and happened to identify unusual repetitive palindromic repeat DNA sequences in the *Escherichia coli* genome (Ishino et al., 1987). It wouldn't be realized until nearly 30 years later the powerful gene editing tool that was accidentally discovered in this study. In 2002, Ruud Jansen and colleagues identified genes associated with these CRISPR sequences which they labeled "Cas" genes. (Jansen et al., 2002). Other studies confirmed that the CRISPR-Cas system

functioned as one bacterial defense system. Following the discovery of the Cas9 protein which essentially acts as “genetic scissors”, researchers such as Jennifer Doudna and Emmanuelle Charpentier discovered that the CRISPR-Cas9 system could be programmed to target specific DNA sequences and could introduce double-stranded breaks (DSBs) to edit the sequence in simpler organisms such as bacteria (Jinek et al., 2012). Given this, around 2013, researchers including Feng Zhang and his team at the Broad Institute tested the system on mammalian cells and found the results to be successful. Starting as a defense mechanism for bacteriophages, CRISPR-Cas9 has now converted into one the most exciting gene editing tools in the world of gene therapy following a combination of meticulous basic research, and innovative translational experiments (Ran et al., 2013).

As demonstrated in Figure 3, the CRISPR-Cas system is composed of a Cas protein (a nuclease) to do the editing, a guide RNA (sgRNA) which locates the Cas protein to the targeted DNA sequence anywhere in the genome, and a PAM sequence. Once the complex has reached its targeted site, the PAM will allow the Cas9 protein to bind to the DNA sequence. Cas9 will then unzip the DNA, find a match with its target RNA, and introduce a DSB in the DNA sequence. Through natural repairing processes like Non-Homologous End Joining (NHEJ) and Homologous Directed Repair (HDR), the DNA will repair itself through the error-prone process using NHEJ (M. Asmamaw et al., 2021). Now a short 20-base pair sequence is enough to locate a nuclease editor to a unique region in the genome to execute gene modification in a precise and efficient manner. These guide RNAs can be designed and tested in a matter of weeks compared to months or years that were once required with early editors.

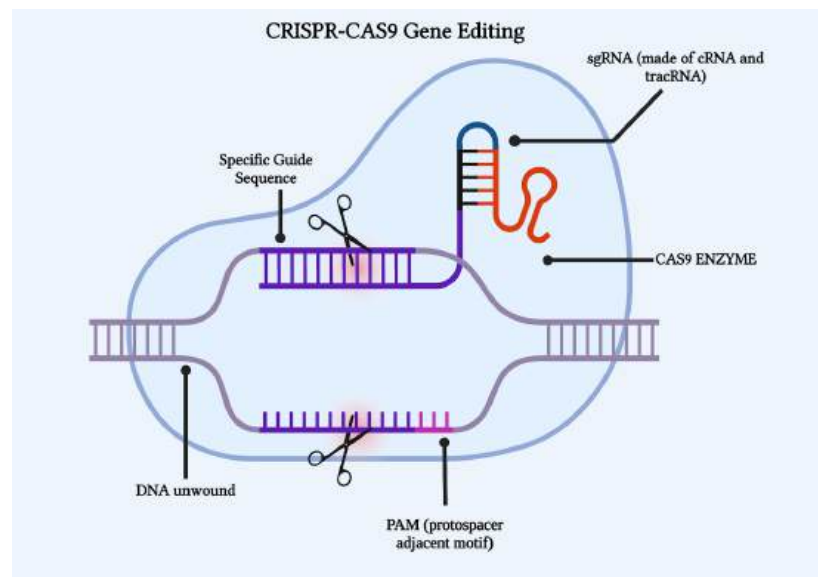


Figure 3: Structure of CRISPR-Cas9 system

The NHEJ pathway is the primary repairing pathway due to its ability to activate throughout all stages of the cell's cycle. This method is the faster repairing process, relying on proteins that help bring the broken ends of the DNA strands closer together and ligase enzymes to stitch the ends. Expectedly, this fast repair process can cause mistakes due to the lack of a template, quickness, and the complexity of end joining. Unlike NHEJ which is active in all stages of the cell cycle, HDR is mainly restricted to phases where DNA is accessible as it relies on a homologous DNA template to fix breaks. Despite its ability to achieve higher precision than NHEJ due to its use of a DNA template, there are still possibilities of an imperfect template. Overall, both repair processes are limited and flawed, making them unpredictable, unreliable, and even risky (Yang et al., 2020). The mechanism of these natural processes are summarized and compared in Figure 4. These processes can lead to severe side effects as they are not guaranteed to repair the genes perfectly. Minor insertions and deletions in a DNA sequence can lead to major mutations, chromosomal rearrangements, and overall, genomic instability. As a result, new innovative approaches like base editing and prime editing, which are based on CRISPR-Cas9 systems, have found ways to precisely edit DNA sequences without introducing DSBs allowing for a much safer and predictable outcome. These techniques nick single strands of DNA rather than both strands of the DNA and can either change single base pair combinations or replace DNA sequences up to 44 base pairs incredibly precisely. As a result, the mistakes from natural DNA repairing processes can be avoided (Newby and Liu, 2021).

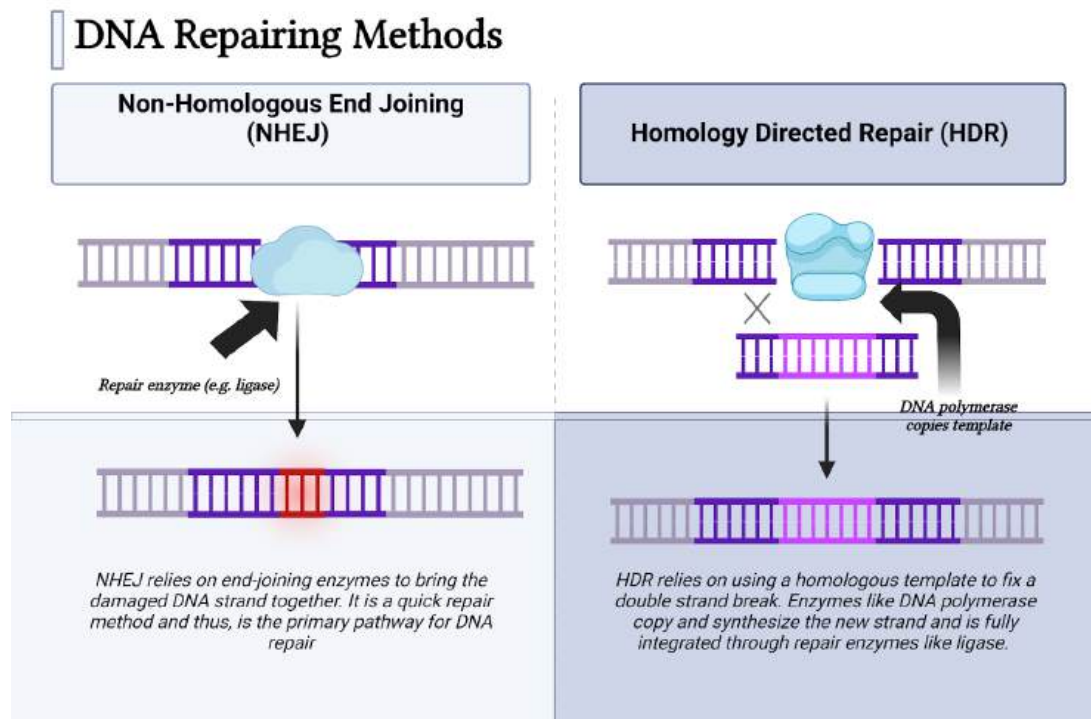


Figure 4: DNA Repairing Methods

Gene therapy using *ex vivo* CRISPR

Following advances in research and translational experiments in animals, CRISPR-Cas9 evolved from a research tool to a therapeutic application. The first clinical trial of CRISPR-Cas9, performed by You Lu and colleagues, treated non-small-cell lung cancer by editing T-cells *ex vivo*. The treatment used two sgRNAs to target the PD-1 gene—a gene exploited by cancer cells to evade an immune response—and a Cas9-encoding plasmid DNA, both electroporated into the cell. The study highlighted the clinical results of patient B-01 who did not experience a progression of the disease, but rather, remained stable for 76 weeks, a prolonged T-cell expression with high TCR diversity suggesting the successes of the treatment that could arrive in the future (Lu et al., 2020). Overall, the study marked the first step toward the potential therapeutic applications CRISPR-Cas9 offers.

To use CRISPR to treat HIV, previous research has suggested that the disruption of the CCR5 gene, a crucial receptor for the entry of HIV cells, could make cells resistant to HIV-1. Thus, a study used CRISPR-Cas9 to edit the CCR5 gene in hematopoietic stem cells and progenitor cells to treat a 27-year-old man diagnosed with HIV and acute lymphoblastic leukemia. Healthy CD34+ cells from a donor were edited to disrupt the CCR5 gene as well as unedited CD34+ genes were infused into the patient to support engraftment and immune recovery. Cells had successfully engrafted, and the patient's bone marrow did not demonstrate any detectable signs of leukemia (Xu et al., 2019).

Another key study using CRISPR focused on treating Sickle Cell Disease and β -thalassemia, monogenic diseases severely hindering the transportation of oxygen in the bloodstream. The team of researchers used CRISPR-Cas9 to edit the BCL11A transcription factor, a series of proteins that represses γ -globin expression and fetal hemoglobin in erythroid cells, in healthy hematopoietic stem cells to enhance the γ -globin and fetal hemoglobin expression. The results of the study were impressive, as a patient with β -thalassemia experienced a rise in fetal hemoglobin in erythrocytes from 10.1% to 99.7% within 6 months (H. Frangoul et al., 2021).

Treatments using CRISPR for the retina have also proved to be successful. Editas Medicine developed a treatment called EDIT-101 for Leber congenital amaurosis, a genetic disorder causing severe vision loss and potential blindness. Through the use of AAVs, Editas Medicine delivered a CRISPR-Cas9 complex targeting the dysfunctional CEP290 gene within the patient's retinal cells to restore normal function. With no serious adverse effects, the treatment was fairly successful with around 6 out of 12 patients reporting meaningful improvement in their vision (Pierce et al., 2024).

Gene therapy using *in vivo* CRISPR

More recent clinical trials of CRISPR-Cas9 have deviated from *ex vivo* approaches and delved into *in vivo* approaches via targeting delivery agents such as viral vectors and nanoparticles. Now, researchers and scientists are combining their long history of improving DDS with their advancements in gene editing to revolutionize the treatment of genetic conditions. An innovative study developed a unique nanocapsule to encapsulate the

CRISPR-Cas9 complex to treat glioblastoma in mice. The nanocapsule consisted of a polymeric shell that would break down when in contact with glutathione, a protein overly expressed in tumor cells. Functionalized with an angiopep2 peptide targeting the LRP-1 receptor, the complex crossed the blood-brain barrier and encapsulated the CRISPR-Cas9 complex with a guiding RNA to direct the complex to the PLK-1 gene, which is heavily associated with cancer. The nanocapsule showed remarkable results with a large reduction in the PLK-1 gene and cell death among glioblastoma cells (Zou et al., 2022).

In addition, Intellia Therapeutics developed a highly successful *in vivo* CRISPR-Cas9 therapy called NTLA-2001 for treating Transthyretin Amyloidosis, a fatal disease caused by the accumulation of the misfolded transthyretin proteins in tissues typically in the heart and nerves commonly leading to cardiomyopathy. NTLA-2001 is composed of the CRISPR-Cas9 system guided by a sgRNA targeting the specific TTR sequence to cut production and is delivered by the apolipoprotein E-coated lipid nanoparticles which are easily recognized by liver cell receptors facilitating endocytosis. With this innovative targeting system, an average TTR reduction concentration of 87% was demonstrated with minimal off-targeting results and mild adverse events. (Gillmore et al., 2021). Another drug in the Intellia Therapeutics pipeline is for the treatment of hereditary angioedema, a disease-causing unpredictable and severe swelling associated with severe complications like obstructed airways, and anaphylactic shocks. Similar to NTLA-2001, NTLA-2002 uses lipid nanoparticles to deliver the CRISPR-Cas9 mRNA targeting KLKB1 to hepatocytes. This reduces plasma prekallikrein, a protein that matures into the kallikrein enzyme which boosts the production of bradykinin, a peptide that can cause vasodilation, and increased vascular permeability contributing to the symptoms of angioedema. In clinical studies, NTLA-2002 treatment demonstrated a near-complete reduction in both kallikrein levels and angioedema attacks. The study demonstrates the effectiveness of NTLA-2002 as well as the endless possibilities for CRISPR-Cas9 systems simply by changing the sgRNA (Longhurst et al., 2024).

Beyond Double-Stranded Breaks - Base Editors and Prime Editors

Despite all of the successes in CRISPR-Cas9 studies, its nuclease-based approach to editing genes and introducing double-stranded breaks can pose implications. Natural repairing processes like nonhomologous end joining and microhomology-mediated end joining are triggered in response to the double-stranded breaks. However, these repairing processes are fast and error-prone, often leading to indels at the break site which can disrupt gene function, cause unstable chromosomal rearrangement, and mutations that can lead to cancer. As a result, the field of gene editing has evolved from traditional nuclease-based editing to advanced genomic editing systems such as base editing and prime editing which use single-stranded breaks rather than double-stranded breaks. Through a series of specific mutations, both base and prime editing convert the Cas9 enzyme into a nickase, a variant that only nicks a single strand of the DNA rather than both. These methods offer higher precision and safety making them the future of gene editing.

Base editing is composed of the modified Cas9 nickase, a deaminase enzyme responsible for changing base sequences in the DNA, and fused with a single guiding RNA for the Cas9-deaminase complex to the targeted DNA sequence as demonstrated in Figure 5. Once the sgRNA has guided the complex to the specific DNA site, the Cas9 nickase will unwind the double helix structure of the DNA and search for a perfect match between the sgRNA and the targeted DNA sequence. Once this match has been identified, the Cas9 nickase will nick a single strand of the DNA while the deaminase enzyme will change the base. Most base editors will take an intermediate step before changing a specific base. When editing Cytosine to Thymine, the deaminase enzyme will edit the Cytosine to Uracil, which is then recognized as an error by natural DNA repairing processes and will change the base to Thymine. A similar process occurs with Adenine and Guanine base pairs however, instead of Uracil, the intermediate step is Inosine. Thus, a highly precise base editing technology is born (Newby and Liu, 2021).

Contrasting traditional gene therapies which separate a sgRNA and a donor RNA template, a prime editing complex uses a prime editing guide RNA (pegRNA) which combines both, a complex demonstrated in Figure 5. In addition to the pegRNA, the complex contains the Cas9 nickase, a reverse transcriptase enzyme, and lastly, a primer binding site. Once the complex has been administered into the body, the pegRNA finds the targeted sequence and through the primer binding site, the Cas9 nickase will be able to bind to the sequence and begin its job. After introducing a single-stranded break, the reverse transcriptase will use the pegRNA as a template, initiating reverse transcription and binding the appropriate nucleotides to the nicked sequence. However, because only one strand of the DNA is edited and the other is not, a mismatch occurs in the overall structure of the double helix. Repair pathways recognize this and use the prime-edited strand as a template to create a complementary sequence to replace the old, unedited sequence. In essence, prime editing is the same fundamental idea as other nuclease-based editing technologies however, it breaks the steps of genomic editing into smaller, safer, and more precise steps, as seen by the subsequent single-stranded breaks (Newby and Liu, 2021). Both novel editing technologies are summarized in Figure 5.

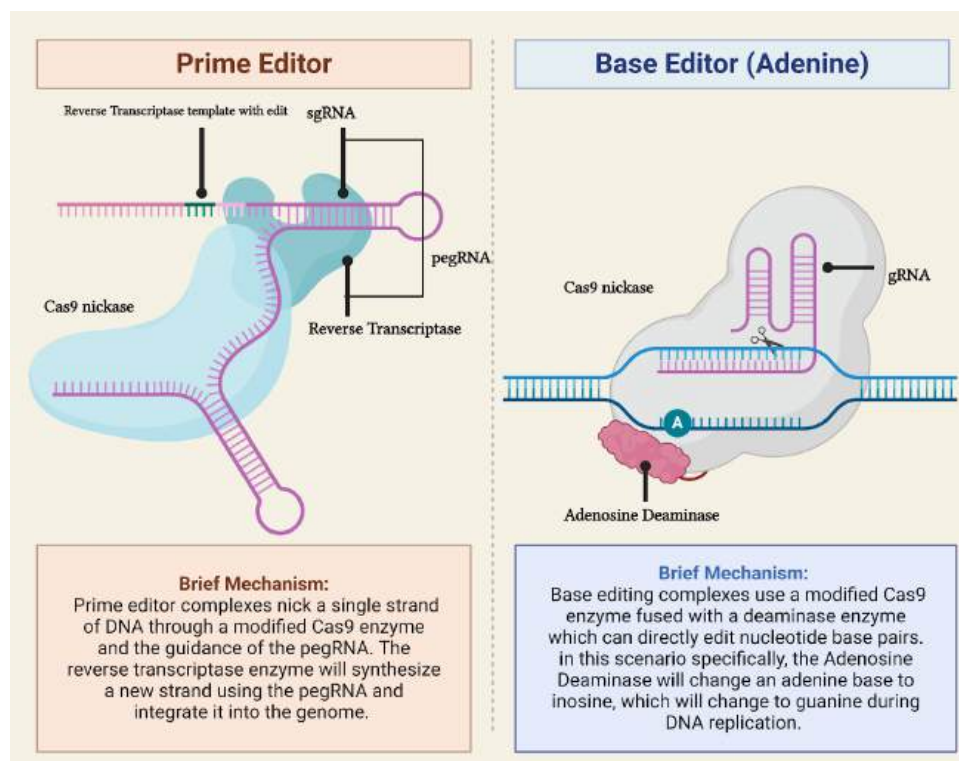


Figure 5: Prime Editing and Base Editing

Both base editing and prime editing have been successful in a multitude of clinical studies. One treatment using base editing focused on CD3 δ SCID, a severe genetic disorder caused by mutations in the CD3 δ gene, responsible for encoding the CD3 δ chain which is vital for functional thymopoiesis. An adenine base editing system was developed to treat the immunodeficiency consisting of a guide RNA, and an mRNA encoding the Cas9 nickase and adenine deaminase enzyme. After extracting cells from CD3 δ SCID patients, electroporation was performed to edit the unhealthy cells and transplanted into mice to measure their efficacy. The study showed successful restoration of T cell function with one variant achieving an almost perfect correction rate and reaching the point where the cells were comparable to healthy cells. Overall, the T cells matured into fully functional T cells with long-term engraftment in mice with minimal off-target results demonstrating the capabilities of these new Cas9 variants (McAuley et al., 2023). Another base editing study treated Duchenne muscular dystrophy (DMD), an acute genetic disease caused by mutations in the DMD gene which produces the dystrophin protein, a protein vital for muscle cell stability. To facilitate the translation of CRISPR-based therapies for human DMD patients, researchers developed a humanized mouse model of DMD replacing exons 50 and 51 of the DMD gene with human exon sequences. An adenine base editor complex containing a deaminase enzyme and three sgRNAs were delivered to cells via an adeno-associated virus (AAV-ABE2) and demonstrated the most successful results in neonatal mice, almost completely restoring dystrophin expression and improving muscle strength and

function comparable to healthy mice (Lin et al., 2024). Overall, the success of base editing in mice demonstrated much promise for clinical applications.

The first clinical trial of base editing, done in November of 2023 by Verve Therapeutics, shows potential in reducing cholesterol levels. The methodology involves an injection treatment labeled as VERVE-101 to target and suppress the PCSK9 gene in the liver which is directly responsible for the production of low-density lipoprotein cholesterol. This type of cholesterol is commonly associated with fatty plaque buildup and cardiovascular conditions. VERVE-101 is composed of a mRNA encoding the base editor enzyme, a guide RNA engineered to recognize the targeted sequence in the PCSK9 gene and delivered through the lipid nanoparticles designed with liver tropism. The treatment, aiming for a one-time dosage cure to conditions associated with high cholesterol levels, showed relatively successful results demonstrating a 55% reduction in LDL cholesterol levels and a high expression reduction in the PCSK9 gene. However, out of the ten patients, two of them had experienced heart attacks following the VERVE-101 treatment with one of them passing away. While having relatively successful results, meticulous study must be done on the safety of the treatment in the following years (M. Naddaf, 2023)

Similar to base editing, prime editing has been successful in the clinic. To treat cystic fibrosis in humans, a study used a prime editing complex composed of a modified Cas9 nuclease (SpCas9 H840A), a reverse transcriptase, and a pegRNA to edit the mutation. Cystic fibrosis is a genetic disorder caused by the F508del mutation which causes a disruptive buildup of mucus typically in the lungs and digestive system. The mutation causes the cystic fibrosis transmembrane conductance regulator protein to misfold, a protein regulating the balance of salt and water on the surfaces of organs to keep mucus at a thin consistency. Initial attempts were essentially futile however with advances in pegRNAs, an MLH1dn protein to reduce mismatch repairs, and advances in prime editing complexes led to an increase in genomic corrections, a 140-fold improvement in bronchial epithelial cells, and a 26-fold improvement in embryonic kidney cells (Sousa et al., 2024). The results of the study highlight the importance of optimization in gene editing techniques like prime editing, further solidifying the potential of prime editing.

Concerns and Challenges

Overall, the transition from *ex vivo* gene therapies to *in vivo* gene therapies has proven to be incredibly successful and promising. The combination of novel gene editing technologies along with carefully engineered delivery modalities has shown therapeutic success in both animals and humans. However, despite the successes and potential of delivery modalities such as LNPs, PNP, and VLP, many challenges and limitations still arise. Pfizer's and Moderna's LNP vaccine remains one of the only clinically approved LNP treatments by the FDA. The majority of these concerns lie in the potential immunogenicity associated with the delivery modality. Namely, the COVID-19 vaccines have been positively associated with adverse effects such as autoimmune diseases and allergies. Myocarditis, the inflammation of the heart, and thrombocytopenia, the state of having low blood platelets, are among the conditions that have

been increasingly associated with the COVID-19 vaccines (Lee et al., 2023). Furthermore, other studies concerning delivery modalities have found that when studying cancer therapies in specific, nanocarriers such as LNPs, PNPs, Dendrimers, and Iron Oxide NPs caused significant inflammatory responses, and cytotoxicity once administered into the body. Due to the small sizes of all nano-delivery systems, there exists an increased risk of unforeseen and potentially dangerous biological interactions as they can easily pass through cell membranes, and engage with cell receptors. Additionally, the transition from lab to clinic demonstrates high difficulty as they must pass meticulous safety concerns, slowing down the process of innovating new and effective delivery systems. (M Chehelgerdi et al., 2023).

Similar to other nanoparticles, VLPs are also struggling to get past the preclinical phase of their trials. Few therapies have been approved due to difficulties in and outside the lab. The production of VLPs has been relatively inconsistent as their reliance on the self-assembly and folding of proteins can often go wrong. This hinders not only the functionality of the VLPs but also makes the production process ineffective and expensive. (Gupta et al., 2023).

However, the concerns of gene therapies do not solely depend on the delivery modality. While CRISPR-Cas9 systems and their variants, base editing, and prime editing, have revolutionized the world of genomic editing, it is still among one of the most controversial topics of modern-day science. As has always been, the leading concern regarding genomic editing is its precision. Off-target effects can unintentionally cause severe mutations, an example being unintentionally editing a tumor-suppressing gene to be deactivated causing the proliferation of cancer cells to be extremely challenging to fight (Newby and Liu, 2021). Other CRISPR studies have shown severe adverse effects. A CAR-T treatment for relapsed T-cell acute lymphoblastic leukemia using base editing to prevent fratricide and host vs graft disease led to patients undergoing bacterial infections, respiratory issues, central nervous system disease, and death (R. Chiesa et al., 2023). The first clinical trial of base editing attempting to reduce LDL cholesterol levels showed promising results however did lead to 2 patients experiencing heart attacks with one of them passing away (M. Naddaf, 2023). More recent concerns regarding base and prime editing center around genotoxic effects. Studies have shown that prime editing and base editing, despite their precision, are still prone to causing mutations, can trigger immune responses and cell apoptosis via the p53 pathway unintentionally, and can even cause double-stranded breaks defeating the principle purpose of their function (M. Fiumara et al., 2024).

Ethical concerns also arise from these new advancements. The efficacy of such gene therapies has brought many debates regarding eugenics and using genetic modification outside of medical purposes. The idea of a genetically modified human raises various questions on how it would affect regulation, social and equity issues, and overall, the fundamental concepts of human identity. Further concerns lay around the financial aspect of gene editing, an ethical issue commonly debated throughout the history of medicine. Many concerns are that even if gene editing therapies were to overcome their flaws, they would be exclusive to the wealthy and elite due to the cost of development, limited availability, and overall complexity of the treatment (F. Ayanoglu et al., 2020).

Further research is required in order to investigate these safety and social issues. The complexity of nanocarriers makes this process difficult and laborious, yet it is necessary as the use of nanotechnology in medicine is increasing exponentially. Interactions between designed modalities and cellular environments must be carefully examined in order to translate lab results into the clinic. Similarly, advancements in CRISPR in its precision and efficacy must also take the same route. Producing precise, biocompatible, and cost-effective gene editing technologies are among the leading factors in advancing the scientific frontier. While these difficulties among all delivery modalities are strictly limiting, they should not be viewed as flawed technology. As all innovations need, these incredibly new technologies need time. The process of developing an effective treatment and achieving FDA approval typically takes 10-15 years. A laborious process consisting of years of pre-clinical research, years of clinical trials split up into three phases, and years of attempted approval by the FDA. The successes of many of these treatments demonstrated in the clinic guarantee that in the coming years, these technologies will evolve into a versatile fully-fledged treatment for genetic conditions

Future Directions

Here we have reviewed the current implementation of gene therapy across a range of diseases. The journey from discovery to preclinical efficacy often spans over five years unless bolstered by prior academic groundwork. To accelerate this, predictive AI tools are increasingly used to optimize preclinical studies, transforming therapeutic delivery with a focus on efficiency and efficacy. AlphaFold was the first company to use AI to solve fundamental biological concepts. By using an advanced deep learning model, they have predicted over 200 million protein structures to understand their biological function. Traditional methods like X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy are costly and time-intensive, with 90% of protein structures still unsolved. AlphaFold, developed by DeepMind, addresses this using AI to predict protein structures based on evolutionary relationships and folding principles. In CASP14, AlphaFold achieved remarkable accuracy, matching experimental results for most targets (Kryshtafovych et al., 2021). This breakthrough accelerates drug development, offering insights into unknown genes and enabling the design of stable, precise, and non-immunogenic nanoparticles for therapeutic delivery.

Nanite Bio is another innovative company applying AI to drug delivery by developing a non-viral gene delivery system using PNPs to deliver genetic material to specific tissues. The technology uses an advanced machine learning algorithm known as SAYER which can explore a vast array of polymer combinations and identify the most optimal combination. SAYER accelerates development by rapidly refining and interpreting nanoparticle designs, a process traditionally slow and labor-intensive. Furthermore, the non-viral approach reduces immunogenicity, potentially preventing severe adverse effects. Currently, Nanite is partnering with SalioGen Gene Coding to create a PNP targeting the lungs to develop a permanent non-viral treatment for cystic fibrosis. The nanoparticle would utilize a bioengineered enzyme to integrate the CFTR gene into the lung cells to correct the mutation and prevent the buildup of the

potentially fatal mucus (Nanite, 2024). Overall, AI's use in medicine continues to prove its potential. Similarly, Exscientia has focused on the potential of AI applying it to a drug discovery system. Their advanced AI models analyze large data sets that can simultaneously process and optimize properties of drug candidates such as biocompatibility, stability, potency, and safety, and simulate their interactions with their target. The company uses patient-driven data, analyzing real human tissue to replicate a clinical environment. Currently, Exscientia is working on developing drugs to target specific proteins that are expected to be involved with a variety of conditions such as myeloma, Huntington's Disease, solid tumors, hematological malignancies, and many more. Overall, the delivery and creation of therapeutics show much promise for the future. Through AI's ability to discover drugs in a parallel manner rather than sequentially, efficacy and costs are optimized. Biotech companies such as Alpha Fold, Nanite, and Exscientia reflect the importance of interdisciplinary studies as they can modernize in different ways by implementing cutting-edge technologies.

Currently, off-target effects and safety concerns are among the most discussed issues regarding gene editing. As a result, new innovative approaches like base editing and prime editing are at the forefront of creating more precise, efficient, and cost-effective technologies. However, these new editing approaches are still in development and their therapeutic efficacy remains to be proven. Many steps are being taken in order to improve these next-generation editors. For CRISPR, a variety of variants are being developed to improve specific functions. Mutations introduced in the REC3 domain of CRISPR which helps with targeting have led to variants known as Cas9-HF1 which greatly improves specificity (Skeens et al., 2024). Cas12f, a relative to the Cas9 nuclease, is a much smaller nuclease that can be delivered into cells more easily and is thus being implemented into many genetic treatments (Wu et al., 2024). Similar to Cas9 variants, mutations are being introduced to base editor deaminases to narrow the editing window, for instance, mutations in the rAPOBEC1 cytidine deaminase used in BE3 can reduce the editing window from 5 nucleotides (nts) to as small as 1-2 nts. Consequently, the small editing range will lessen the likelihood of editing undesired genes. Prime editing systems have been using dual pegRNAs to edit larger sequences of DNA and have also been altered to create new variants. Systems known as PE4 and PE5 use small-molecule inhibitors to prevent the mismatch repair of the DNA from acting before the prime editing system has made its desired edits, allowing for a higher success rate (Daliri et al., 2024). Through all of these innovations, AI is starting to be used in order to predict the efficacy of certain variants, facilitating the process of optimizing treatments (Pacesa et al., 2024)

These innovations and adaptations from previously established technologies have shown to be promising. Further research on optimizing each one of the gene therapies must be continued as well as integrating other technologies like the usage of AI modeling and nanoparticles with cell type-specific delivery. More clinical studies for base editing and prime editing will be important to validate these new technologies. Overall, more research on optimization and clinical studies will set up these gene editing technologies to be the future

Works Cited

- An, Meirui, et al. "Engineered Virus-like Particles for Transient Delivery of Prime Editor Ribonucleoprotein Complexes in Vivo." *Nature Biotechnology*, 8 Jan. 2024, pp. 1–12, www.nature.com/articles/s41587-023-02078-y, <https://doi.org/10.1038/s41587-023-02078-y>. Accessed 8 Jan. 2024.
- Asmamaw, Misganaw, and Belay Zawdie. "Mechanism and Applications of CRISPR/Cas-9-Mediated Genome Editing." *Biologics : Targets & Therapy*, vol. 15, no. 1, 21 Aug. 2021, pp. 353–361, www.ncbi.nlm.nih.gov/pmc/articles/PMC8388126/.
- Ayanoglu, Fatma Betul, et al. "Bioethical Issues in Genome Editing by the CRISPR/Cas9 Technology." *Turkish Journal of Biology*, vol. 44, no. 2, 2 Apr. 2020, pp. 110–120, www.ncbi.nlm.nih.gov/pmc/articles/PMC7129066/, <https://doi.org/10.3906/biy-1912-52>.
- Bangham, A D, et al. "Diffusion of Univalent Ions across the Lamellae of Swollen Phospholipids." *Journal of Molecular Biology*, vol. 13, no. 1, 1965, pp. 238–52, www.ncbi.nlm.nih.gov/pubmed/5859039, [https://doi.org/10.1016/s0022-2836\(65\)80093-6](https://doi.org/10.1016/s0022-2836(65)80093-6).
- Barenholz, Yechezkel (Chezy). "Doxil® — the First FDA-Approved Nano-Drug: Lessons Learned." *Journal of Controlled Release*, vol. 160, no. 2, June 2012, pp. 117–134, <https://doi.org/10.1016/j.jconrel.2012.03.020>.
- Bhardwaj, Harish, and Rajendra Kumar Jangde. "Current Updated Review on Preparation of Polymeric Nanoparticles for Drug Delivery and Biomedical Applications." *Next Nanotechnology*, vol. 2, 1 June 2023, pp. 100013–100013, <https://doi.org/10.1016/j.nxnano.2023.100013>. Accessed 16 Mar. 2024.
- Chehelgerdi, Mohammad, et al. "Progressing Nanotechnology to Improve Targeted Cancer Treatment: Overcoming Hurdles in Its Clinical Implementation." *Molecular Cancer*, vol. 22, no. 1, 9 Oct. 2023, www.ncbi.nlm.nih.gov/pmc/articles/PMC10561438/, <https://doi.org/10.1186/s12943-023-01865-0>.
- Chiesa, Robert, et al. "Base-Edited CAR7 T Cells for Relapsed T-Cell Acute Lymphoblastic Leukemia." *The New England Journal of Medicine*, vol. 389, no. 10, 14 June 2023, <https://doi.org/10.1056/nejmoa2300709>.
- Doerfler, Walter. "Adenoviruses." *Nih.gov*, University of Texas Medical Branch at Galveston, 2010, www.ncbi.nlm.nih.gov/books/NBK8503/.
- Fiumara, Martina, et al. "Genotoxic Effects of Base and Prime Editing in Human Hematopoietic Stem Cells." *Nature Biotechnology*, 7 Sept. 2023, pp. 1–15, www.nature.com/articles/s41587-023-01915-4, <https://doi.org/10.1038/s41587-023-01915-4>.
- Frangoul, Haydar, et al. "CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia." *New England Journal of Medicine*, vol. 384, no. 3, 5 Dec. 2020, www.nejm.org/doi/full/10.1056/NEJMoa2031054, <https://doi.org/10.1056/nejmoa2031054>.

- Gabizon, Alberto A. “Pegylated Liposomal Doxorubicin: Metamorphosis of an Old Drug into a New Form of Chemotherapy.” *Cancer Investigation*, vol. 19, no. 4, Jan. 2001, pp. 424–436, <https://doi.org/10.1081/cnv-100103136>.
- Gaj, Thomas, et al. “ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering.” *Trends in Biotechnology*, vol. 31, no. 7, July 2013, pp. 397–405, www.ncbi.nlm.nih.gov/pmc/articles/PMC3694601/, <https://doi.org/10.1016/j.tibtech.2013.04.004>.
- Gillmore, Julian D., et al. “CRISPR-Cas9 in Vivo Gene Editing for Transthyretin Amyloidosis.” *New England Journal of Medicine*, vol. 385, no. 6, 26 June 2021, www.nejm.org/doi/full/10.1056/NEJMoa2107454, <https://doi.org/10.1056/nejmoa2107454>.
- Gregoriadis, Gregory, and Brenda E. Ryman. “Fate of Protein-Containing Liposomes Injected into Rats. An Approach to the Treatment of Storage Diseases.” *European Journal of Biochemistry*, vol. 24, no. 3, Jan. 1972, pp. 485–491, <https://doi.org/10.1111/j.1432-1033.1972.tb19710.x>. Accessed 22 June 2021.
- Gupta, Reeshu, et al. “Platforms, Advances, and Technical Challenges in Virus-like Particles-Based Vaccines.” *Frontiers in Immunology*, vol. 14, 9 Feb. 2023, <https://doi.org/10.3389/fimmu.2023.1123805>.
- Ishino, Y, et al. “Nucleotide Sequence of the Iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in Escherichia Coli, and Identification of the Gene Product.” *Journal of Bacteriology*, vol. 169, no. 12, 1 Dec. 1987, pp. 5429–5433, www.ncbi.nlm.nih.gov/pmc/articles/PMC213968/.
- Jansen, Ruud., et al. “Identification of Genes That Are Associated with DNA Repeats in Prokaryotes.” *Molecular Microbiology*, vol. 43, no. 6, Mar. 2002, pp. 1565–1575, <https://doi.org/10.1046/j.1365-2958.2002.02839.x>.
- Jinek, M., et al. “A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.” *Science*, vol. 337, no. 6096, 28 June 2012, pp. 816–821, <https://doi.org/10.1126/science.1225829>.
- Karim Daliri, et al. “Prime Editing and DNA Repair System: Balancing Efficiency with Safety.” *Cells*, vol. 13, no. 10, 17 May 2024, pp. 858–858, pmc.ncbi.nlm.nih.gov/articles/PMC11120019/, <https://doi.org/10.3390/cells13100858>. Accessed 18 Nov. 2024.
- Kryshtafovych, Andriy, et al. “Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XIV.” *Proteins: Structure, Function, and Bioinformatics*, vol. 89, no. 12, 7 Oct. 2021, pp. 1607–1617, <https://doi.org/10.1002/prot.26237>.
- Lee, Yeji, et al. “Immunogenicity of Lipid Nanoparticles and Its Impact on the Efficacy of mRNA Vaccines and Therapeutics.” *Experimental & Molecular Medicine*, 2 Oct. 2023, pp. 1–12, www.nature.com/articles/s12276-023-01086-x, <https://doi.org/10.1038/s12276-023-01086-x>.

- Lin, Jiajia, et al. “Adenine Base Editing-Mediated Exon Skipping Restores Dystrophin in Humanized Duchenne Mouse Model.” *Nature Communications*, vol. 15, no. 1, 15 July 2024, <https://doi.org/10.1038/s41467-024-50340-x>. Accessed 13 Sept. 2024.
- Longhurst, Hilary J, et al. “CRISPR-Cas9 in Vivo Gene Editing of *KLKB1* for Hereditary Angioedema.” *The New England Journal of Medicine*, vol. 390, no. 5, 1 Feb. 2024, pp. 432–441, <https://doi.org/10.1056/nejmoa2309149>.
- Lu, You, et al. “Safety and Feasibility of CRISPR-Edited T Cells in Patients with Refractory Non-Small-Cell Lung Cancer.” *Nature Medicine*, vol. 26, no. 5, 27 Apr. 2020, pp. 732–740, www.nature.com/articles/s41591-020-0840-5, <https://doi.org/10.1038/s41591-020-0840-5>.
- McAuley, Grace E., et al. “Human T Cell Generation Is Restored in CD3 δ Severe Combined Immunodeficiency through Adenine Base Editing.” *Cell*, Mar. 2023, <https://doi.org/10.1016/j.cell.2023.02.027>.
- Mehta, Meenu, et al. “Lipid-Based Nanoparticles for Drug/Gene Delivery: An Overview of the Production Techniques and Difficulties Encountered in Their Industrial Development.” *ACS Materials Science Au*, vol. 3, no. 6, 21 Aug. 2023, <https://doi.org/10.1021/acsmaterialsau.3c00032>.
- Mejía-Méndez, Jorge L., et al. “Virus-like Particles: Fundamentals and Biomedical Applications.” *International Journal of Molecular Sciences*, vol. 23, no. 15, 2 Aug. 2022, p. 8579, <https://doi.org/10.3390/ijms23158579>.
- Milone, Michael C., and Una O’Doherty. “Clinical Use of Lentiviral Vectors.” *Leukemia*, vol. 32, no. 7, 22 Mar. 2018, pp. 1529–1541, www.nature.com/articles/s41375-018-0106-0, <https://doi.org/10.1038/s41375-018-0106-0>.
- Naddaf, Miryam. “First Trial of “Base Editing” in Humans Lowers Cholesterol — but Raises Safety Concerns.” *Nature*, vol. 623, 13 Nov. 2023, www.nature.com/articles/d41586-023-03543-z, <https://doi.org/10.1038/d41586-023-03543-z>.
- Naso, Michael F., et al. “Adeno-Associated Virus (AAV) as a Vector for Gene Therapy.” *BioDrugs*, vol. 31, no. 4, 1 July 2017, pp. 317–334, <https://doi.org/10.1007/s40259-017-0234-5>.
- Newby, Gregory A., and David R. Liu. “In Vivo Somatic Cell Base Editing and Prime Editing.” *Molecular Therapy*, vol. 29, no. 11, Nov. 2021, pp. 3107–3124, <https://doi.org/10.1016/j.ymthe.2021.09.002>. Accessed 10 Nov. 2021.
- Pacesa, Martin, et al. “Past, Present, and Future of CRISPR Genome Editing Technologies.” *Cell*, vol. 187, no. 5, 1 Feb. 2024, pp. 1076–1100, <https://doi.org/10.1016/j.cell.2024.01.042>.
- Park, Haesun, et al. “Evolution of Drug Delivery Systems: From 1950 to 2020 and Beyond.” *Journal of Controlled Release: Official Journal of the Controlled Release Society*, vol. 342, 29 Dec. 2021, pp. 53–65, pubmed.ncbi.nlm.nih.gov/34971694/, <https://doi.org/10.1016/j.jconrel.2021.12.030>.

- Pierce, Eric A., et al. "Gene Editing for CEP290-Associated Retinal Degeneration." *The New England Journal of Medicine*, 6 May 2024, pubmed.ncbi.nlm.nih.gov/38709228/, <https://doi.org/10.1056/NEJMoa2309915>.
- "Platform." <https://www.nanite.com>, 2024, nanitebio.com/platform. Accessed 18 Nov. 2024.
- Ran, F Ann, et al. "Genome Engineering Using the CRISPR-Cas9 System." *Nature Protocols*, vol. 8, no. 11, 24 Oct. 2013, pp. 2281–2308.
- Skeens, Erin, et al. "High-Fidelity, Hyper-Accurate, and Evolved Mutants Rewire Atomic-Level Communication in CRISPR-Cas9." *Science Advances*, vol. 10, no. 10, 6 Mar. 2024, <https://doi.org/10.1126/sciadv.adl1045>. Accessed 3 Oct. 2024.
- Sousa, Alexander A., et al. "Systematic Optimization of Prime Editing for the Efficient Functional Correction of CFTR F508del in Human Airway Epithelial Cells." *Nature Biomedical Engineering*, 10 July 2024, pp. 1–15, www.nature.com/articles/s41551-024-01233-3, <https://doi.org/10.1038/s41551-024-01233-3>.
- Vargason, Ava M., et al. "The Evolution of Commercial Drug Delivery Technologies." *Nature Biomedical Engineering*, vol. 5, 1 Apr. 2021, pp. 1–17, www.nature.com/articles/s41551-021-00698-w, <https://doi.org/10.1038/s41551-021-00698-w>.
- Willard, Huntington F. *The Human Genome: Foundation for Genomic and Precision Medicine*. Academic Press, 2 Dec. 2016, pp. 1–19, www.sciencedirect.com/science/article/abs/pii/B9780128006818000013.
- Wu, Tong, et al. "An Engineered Hypercompact CRISPR-Cas12f System with Boosted Gene-Editing Activity." *Nature Chemical Biology*, 3 July 2023, www.ncbi.nlm.nih.gov/pmc/articles/PMC10625714/, <https://doi.org/10.1038/s41589-023-01380-9>.
- Xu, Lei, et al. "CRISPR-Edited Stem Cells in a Patient with HIV and Acute Lymphocytic Leukemia." *New England Journal of Medicine*, vol. 381, no. 13, 26 Sept. 2019, pp. 1240–1247, www.nejm.org/doi/full/10.1056/NEJMoa1817426, <https://doi.org/10.1056/nejmoa1817426>.
- Yang, Han, et al. "Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks." *International Journal of Molecular Sciences*, vol. 21, no. 18, 4 Sept. 2020, p. 6461, www.ncbi.nlm.nih.gov/pmc/articles/PMC7555059/, <https://doi.org/10.3390/ijms21186461>.
- Zielińska, Aleksandra, et al. "Polymeric Nanoparticles: Production, Characterization, Toxicology and Ecotoxicology." *Molecules*, vol. 25, no. 16, 15 Aug. 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7464532/, <https://doi.org/10.3390/molecules25163731>.
- Zou, Yan, et al. "Blood-Brain Barrier–Penetrating Single CRISPR-Cas9 Nanocapsules for Effective and Safe Glioblastoma Gene Therapy." *Science Advances*, vol. 8, no. 16, 22 Apr. 2022, <https://doi.org/10.1126/sciadv.abm8011>.

Impacts of Global Events on the Consumer Economy of the UAE By Arnav Malhotra

Introduction

In recent years, global events such as the COVID-19 pandemic and the conflict between Russia and Ukraine have directly impacted a multitude of consumer-driven sectors in leading economies around the world. Despite not being an epicenter for any of these global issues, the UAE has experienced firsthand externalities from the chain-reactions incurred by these impacts. This paper looks to explore the severe reduction in consumer spending during these periods of economic instability across five key segments: real estate, education, airlines, hotels, and supermarkets.

The objective of this thesis report is to analyse how extreme global events have altered consumer preferences, spending patterns, and purchasing power in the UAE. Understanding these shifts in consumer behaviour will help in analysing broader economic trends and helps us realise what the effects of these events truly are. This thesis report will also help identify and analyse how different segments in the UAE (for example, airlines, education, real estate, and retail) within the UAE have been affected by global events. Such analyses are vital to refining our understanding of the different ways in which different segments react—and are affected by—events of global significance. My analysis, therefore, seeks to provide a more detailed view of the impacts of recent events on varied sectors of the UAE's consumer economy.

The scope of this report consists of interpretations of data that demonstrates how global events affect: the different segments of the UAE; the consumer economy of the UAE, and the economy of the UAE as a whole. By limiting the focus of this research to the UAE, the aim of this paper is to demonstrate the impacts of global events on a unique national economy within the network of international trade. In many ways, the UAE stands as a microcosm of global trade in the way organizations and labour from the Global North and Global South interact in a relatively small space. As such, the UAE, I will argue, offers us insight into the effects of events of global significance on the wider global economy.

Methodology

This study uses a mixed-methods research design that integrates quantitative and qualitative techniques to give a thorough understanding of how global events have affected the UAE's consumer economy. While the qualitative approach will offer deeper insights into the contextual and perceptual components of these consequences, the quantitative approach will make it possible to assess the impacts as the numbers and figures will provide a clear unbiased view.

Data Collection

The data for analysing the impacts of global events on the UAE consumer economy was accumulated through both primary and secondary sources. Primary data was collected by using a payments company's raw data. After that, the necessary data was extracted and put into a bar

chart. Secondary data included thesis reports from many diverse sources, such as media articles, which provide an authentic setting and insight into broader financial patterns. The integration of these information sources permitted for a thorough investigation of the impacts of global events on consumer behaviour and the changing economic patterns within the UAE.

COVID-19 Pandemic

Qualitative Data Analysis

The COVID-19 pandemic has had a significant impact on worldwide economies. The UAE was no exception to this. A later proposal investigating this effect uncovers noteworthy shifts in customer behaviour and financial patterns all over the UAE.

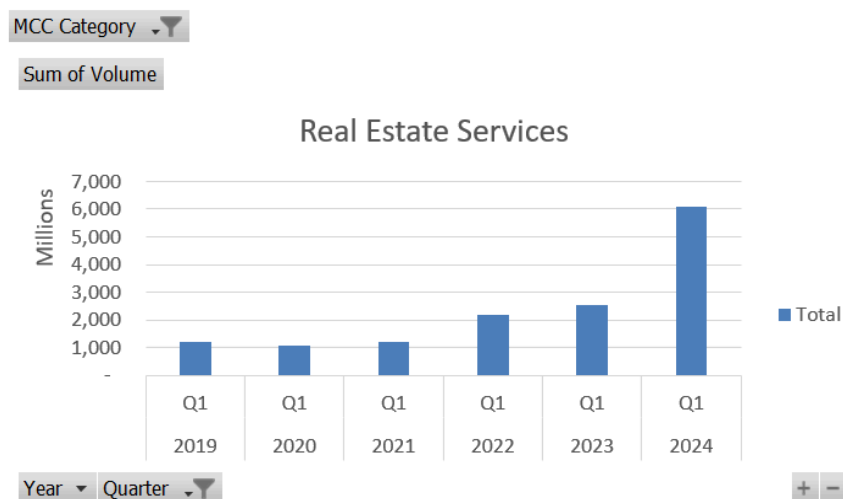
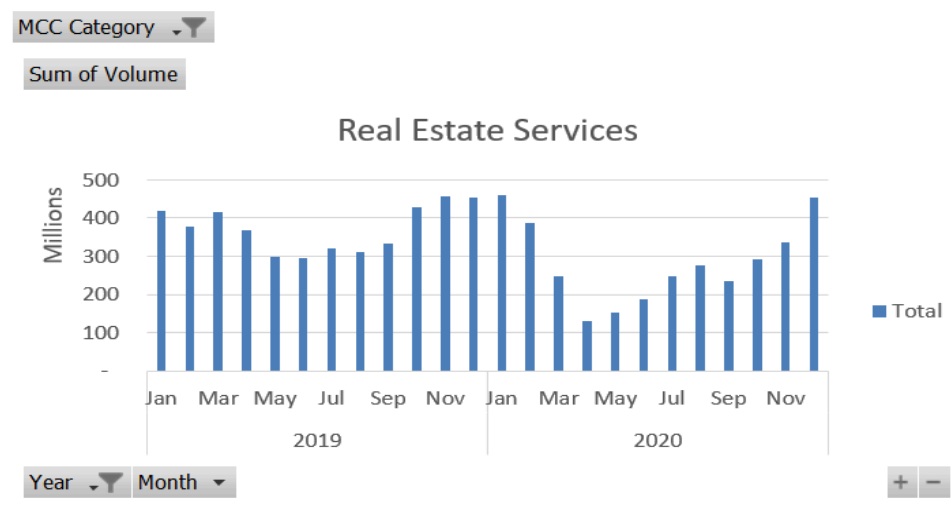
This pandemic disturbed the UAE's consumer economy by influencing different divisions unevenly. Pre-pandemic, the UAE boasted a strong buyer advertisement driven by extravagance investing and tourism. Nevertheless, the onset of COVID-19 led to prompt decreases in these segments. Extravagance retail deals plunged as shoppers prioritized basics and reserve funds over optional investing. The tourism segment confronted serious mishaps, with travel confinements and wellbeing concerns deterring tourists from entering the country.

According to Al Mansoori et al., the COVID-19 pandemic has wreaked havoc on the UAE's economy and the labor market. For example, measures such as full or partial lockdown have resulted in the closing down of companies causing job losses to millions of people. The UAE economy's hardest-hit sectors include real estate, insurance, construction, insurance, travel and tourism, manufacturing, and energy, especially crude oil and natural gas. At the same time, the travel bans and restrictions adversely affected the country's hospitality, tourism, and retail sector. In an attempt to reverse the damage exerted on the economy by the pandemic, the UAE government has taken precise measures, including the exemption of businesses registered onshore from paying corporate taxes and directing financial institution not to take stringent actions against small and medium enterprises during the COVID-19 pandemic

This paints a complex picture of how COVID-19 reshaped the UAE consumer economy. While there were existing patterns moving towards digitalization, the pandemic accelerated their trajectories and highlighted fundamental vulnerabilities in the UAE infrastructure.

Quantitative Data Analysis (Graphs are all for UAE)

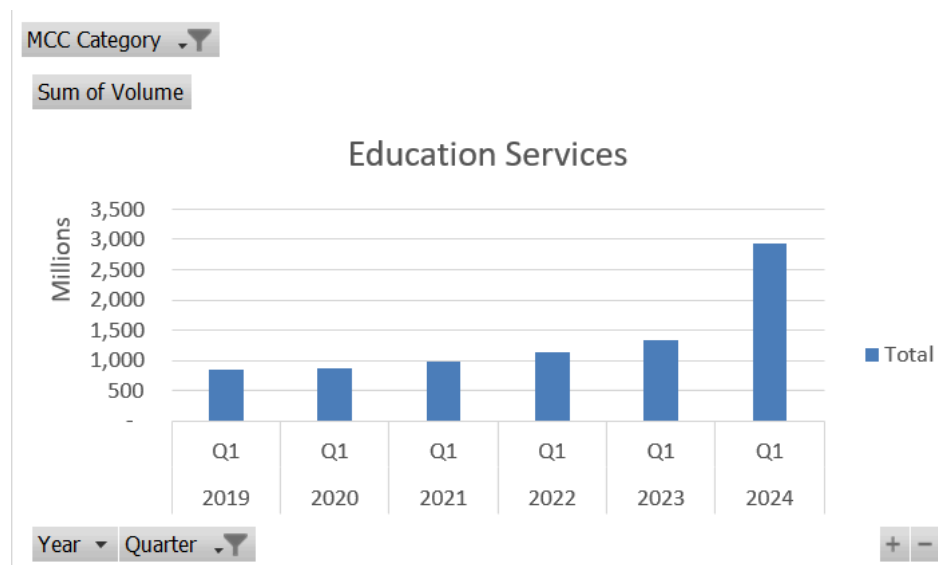
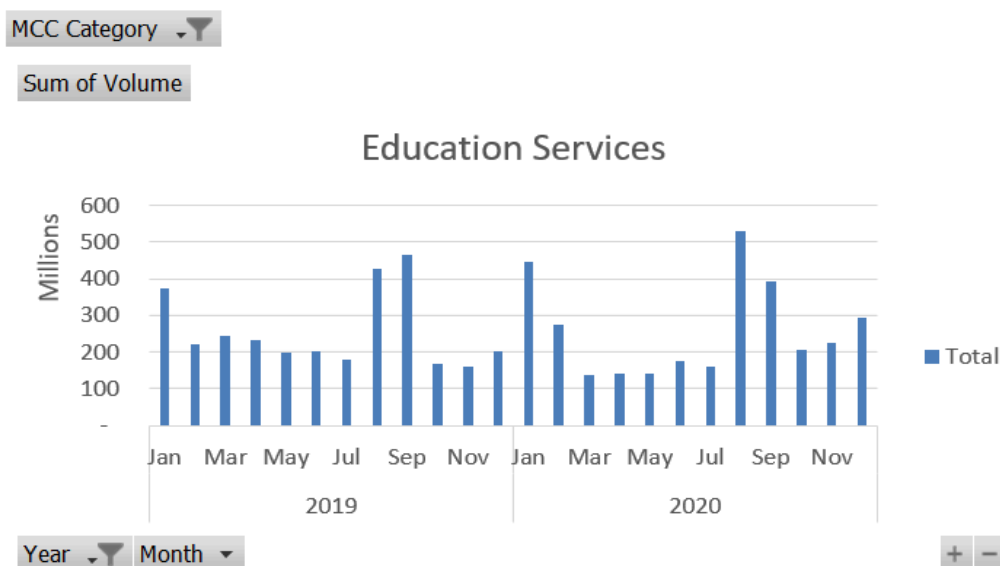
Primary / Pre and Post Covid



Real estate services witnessed a significant decrease in their sum of volume due to COVID-19 as there was a large amount of financial uncertainty. With substantial job losses and business closures, the period's economic instability made it much harder to participate in the real estate market. The pandemic also disturbed the supply chain of the real estate market, which led to many delays and an increase in cost for construction materials as they were harder to transport. Due to social distancing, construction projects slowed down significantly as people were not allowed to be in close contact with each other. Furthermore, as it took longer for projects to be completed, the prices increased significantly, and this discouraged consumers from purchasing properties.

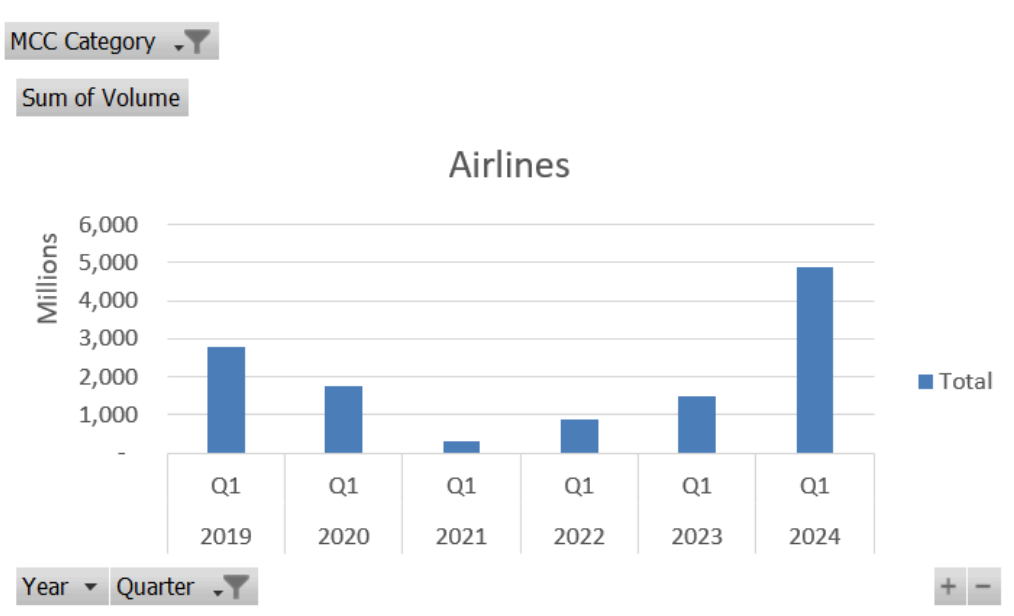
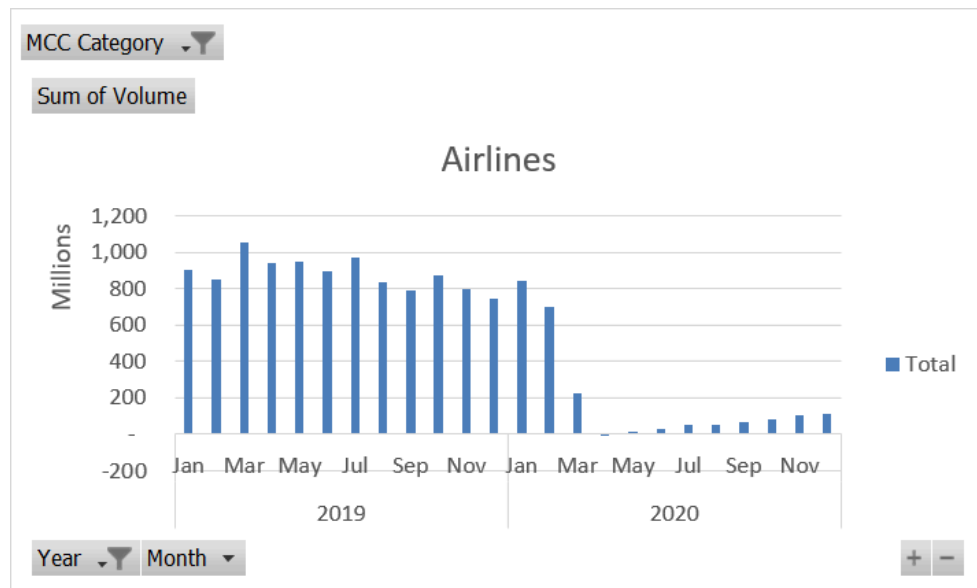
After the end of Q1, however, the sum of volume started to increase again. This could be because many banks in the UAE reduced interest rates to historic lows. This made mortgage loans much more affordable and led to an increase in demand for real estate again. As the pandemic forced people to stay at home, many people felt the need to upgrade their house as they were spending much more time there and they needed more space or better amenities for their

remote work. As there was a lot of uncertainty in Q1 due to the start of the pandemic, many transactions were delayed as no one was sure of what to expect. However, as restrictions eased, more transactions went through and thus the volume of services started to increase again. From April 2019 to April 2020 there was a 65% drop in the sum of volume.

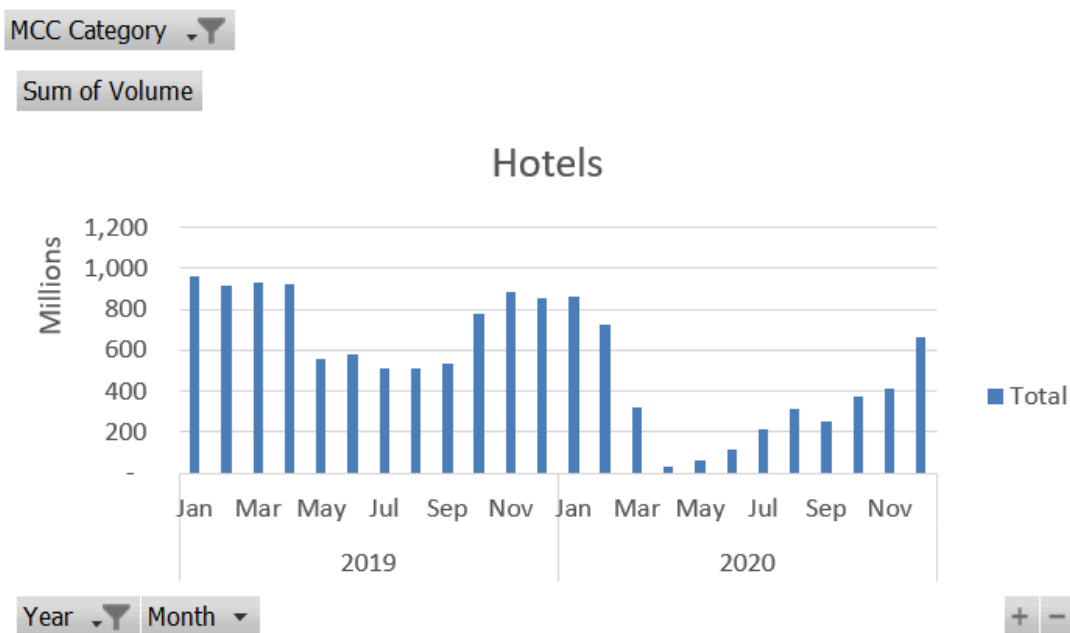


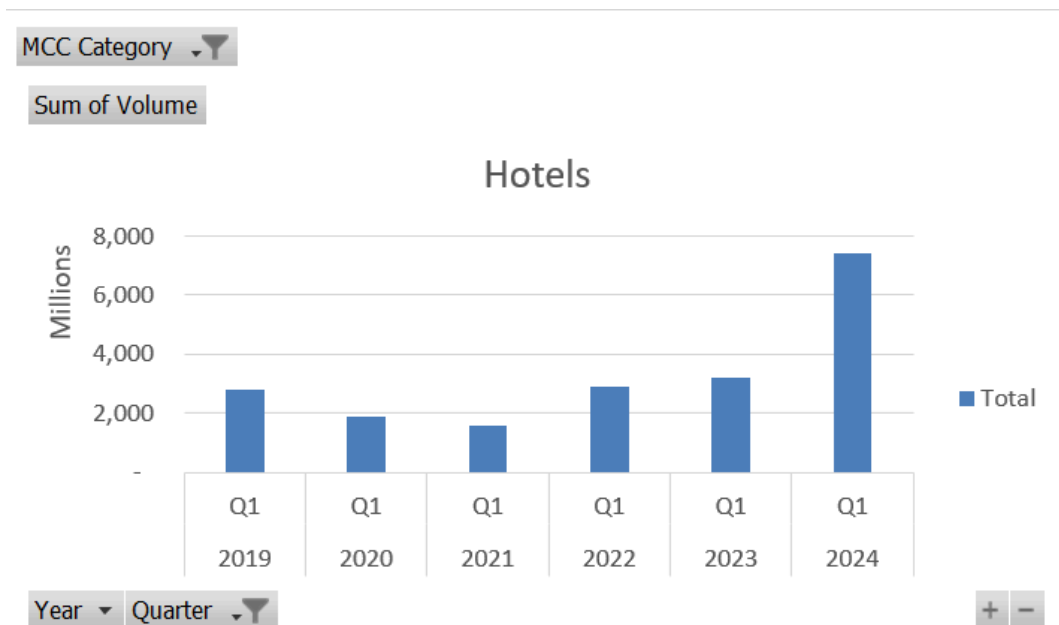
Moving to education services, the sum of volume in this sector increased because of the schools' quick response to the pandemic's disruptions. This indicates that they made the switch to online learning platforms, which led to a sharp rise in demand because students could enrol from anywhere in the world. This was beneficial because some students were stranded in other countries and could participate from anywhere because of the shift to online learning. To facilitate the shift to online learning, the government and numerous educational institutions

supplied extra resources. Investments in technology and other support services were part of this. In addition, some schools offered extra classes to students who couldn't begin their programs at the scheduled time or extended enrolment deadlines. Due to this, education services had a significant increase in revenue. However, it took some time for the pandemic to truly take effect. Due to many job losses, some students could no longer afford to pay for education and had to temporarily take a leave. Some schools and universities also faced many different financial difficulties and due to that they were unable to offer the highest level of education they could, and this led to a decrease in the sum of volume. From April 2019 to April 2020 there was a 40% drop in the sum of volume.



Airlines also faced a significant decrease in their volume as many governments around the world-imposed travel bans to restrict the spread of the virus. The fear of contracting the virus also deterred many individuals from travelling as they wanted to stay safe and not get sick as it was quite lethal in some stages. Due to the global financial recession, a lot of redundancies occurred, and many individuals no longer had jobs. Due to this, people did not have enough disposable income and could not afford any holidays thus reducing revenue for many different airlines. There was also an increase in operational costs for airlines as they had to always sanitize and clean the flight and induce many different health protocols which increased the price of running each flight and therefore reduced their profit. As we can see, in April of 2020, there is a negative sum of volume. This is because many people book their flight in advance, however due to the travel ban being installed, many individuals wanted refunds for the flights they had initially booked. This meant that many airlines had to give money back to the customers and consequently there was a negative sum of volume. From April 2019 to April 2020 there was a 101% drop in the sum of volume.





The enormous drop in domestic and international travel that so severely impacted airlines also affected the hotel industry around 2020 Q1-Q2. This meant that less customers would come to hotels as the number of people travelling decreased significantly. As there was an economic recession, many people had reduced disposable incomes which meant that people could not afford to take any vacations which also led to a decrease in the sum of volume. Hotels also faced increased operational costs as they had to implement health and safety measures, for example different cleaning procedures. However, from 2020 Q1 to Q2, there was an increase in the sum of volume. This is because lockdown restrictions started to ease and there was a gradual increase in travel and accommodation demand. Even though international travel was still heavily discouraged, domestic travel started to pick up. Many people would take small trips around the country or go for a staycation. As a result, more people were staying in hotels. Hotels also implemented many more health protocols to reassure customers that it was safe to stay there. Furthermore, to attract guests, many hotels would have offers or discounts which would pique the interest of many different customers. This all led to an increase in the sum of volume in Q2. From April 2019 to April 2020 there was a 65% drop in the sum of volume.



In contrast to the previously discussed sectors, the supermarket industry not only had a constant sum of volume from 2019 to 2020, it in fact was one of the only sectors to witness an increase. This is because many employees were staying home due to lockdown and social distancing measures and therefore there was an increase in demand for grocery items. Instead of dining out, consumers changed their spending to purchase more ingredients and household supplies from supermarkets. A supermarket holds many necessities like food and water. Therefore, it was deemed essential by the government which meant that it stayed open throughout lockdown and any other restrictions. This meant that there were always customers in the shop stocking up on their food or any other necessity. In the earlier months of the pandemic, many people started to panic buy. They started to hoard many different essential products and one of these products was toilet rolls. This led to an exceedingly high sum of volume for supermarkets as we can see in March 2020 as it is the highest sum of volume in the span of those 2 years. This was driven by economic uncertainty and thoughts of possible supply chain

disruptions. There was also a high demand for e-commerce grocery shopping. As people could order groceries off their phone and get it delivered to their house, this was appealing to many customers and therefore the supermarkets' revenue would increase. From April 2019 to April 2020 there was a 19% increase in the sum of volume.

The War in Ukraine

Qualitative Data Analysis

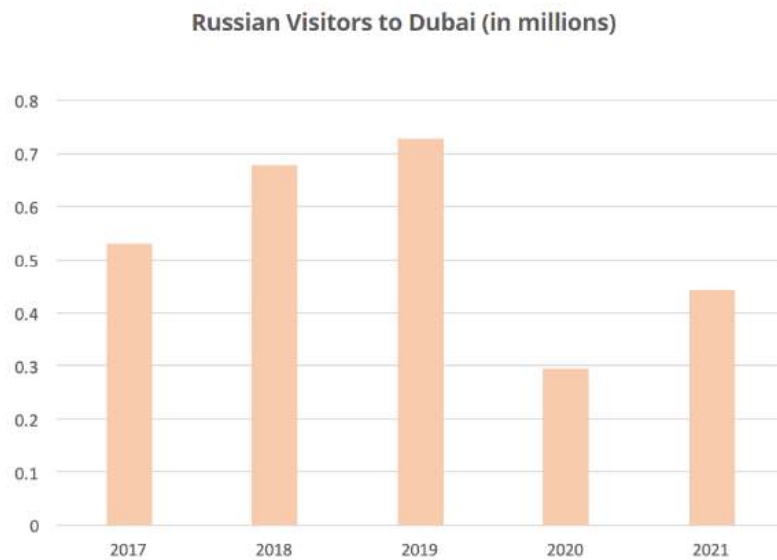
Within weeks of Moscow's full-scale invasion of Ukraine, waves of Russian nationals came to Dubai to reside and invest their money into property, where it would be safeguarded from Western sanctions. Several real estate professionals in the glitzy, tourist-popular emirates told CNBC that Russian parties made up the highest proportion of transactions by nationality for the past year among their client bases. Sajwani estimated Russian nationals comprised about 15% of his customer base. Dubai was already seeing its hottest real estate market in years within the early months of the war. Sales in the sector rose 45% year-on-year in April 2022 and 51% in May, according to the Dubai Land Department.

Following the invasion of Ukraine in February 2022, Dubai experienced a notable increase in Russian nationals relocating to the city. This migration was driven by individuals seeking to invest in real estate, establish businesses, or evade financial restrictions imposed in Europe. Reports indicate that Russian investments in Dubai's real estate market since 2022 have amounted to approximately \$6.3 billion. This surge has contributed to a 124% increase in property prices since 2020, making Dubai a prime destination for Russian capital.

The UAE has refrained from explicitly siding with either Russia or Ukraine in the conflict. It has avoided imposing sanctions on Russia, unlike many Western countries, allowing Russian nationals and businesses to continue operating in the UAE. The UAE has maintained open channels of communication with both Russia and Ukraine. It has supported resolutions at the United Nations condemning the invasion of Ukraine but stopped short of direct criticism or punitive actions against Russia.

In summary, the Russia-Ukraine war has led to increased economic engagement between Russia and the UAE, particularly in trade and real estate. The UAE's neutral stance has bolstered its diplomatic role but has also subjected it to international scrutiny regarding sanctions enforcement and financial transparency.

Quantitative Data Analysis



Source: *Tourism Performance Report, Dubai Tourism, 2017-21*

What the Ukraine Crisis Means for Gulf Economies

The emirate of Dubai in the United Arab Emirates hosts large numbers of Russian tourists annually. In 2021, Dubai was the destination of choice for 6% (600,000 visitors) of tourists from Russia, ranking after Egypt with 10% (1 million) and Turkey with 45% (4.7 million), according to the Russian Association of Tour Operators.

Dubai continues to benefit from the Russia-Ukraine war, Damac chair says

Within weeks of Moscow's full-scale invasion of Ukraine, waves of Russian nationals came to Dubai to reside and invest their money into property, where it would be safeguarded from Western sanctions. Several real estate professionals in the glitzy, tourist-popular emirates told CNBC that Russian parties made up the highest proportion of transactions by nationality for the past year among their client bases. Sajwani estimated Russian nationals comprised about 15% of his customer base. Dubai was already seeing its hottest real estate market in years within the early months of the war. Sales in the sector rose 45% year-on-year in April 2022 and 51% in May, according to the Dubai Land Department.

Overall impact of the Russia and Ukraine war on UAE

The war has led to an increase in global gas and oil prices. As the UAE is a very big player in this market and one of the leading oil exporters, UAE benefited from higher oil revenues. But this could have also led to higher costs for some imported goods and services as prices increased all over the world. Due to the war, many global supply chains have been disturbed, and due to this there are shortages of certain goods in the UAE. The trade routes have

also been disturbed as you could not fly over that region, and this also adds to why the supply chains have been disturbed. As there has been an increase in global inflation, the cost of living for residents in the UAE has significantly increased, leading to higher prices for goods and services and individuals and businesses need to make a higher revenue. UAE has felt the need to diversify its economic portfolio by expanding to other regions which aren't oil and gas. For example, the UAE has increased the amount of money it is spending on technology, finance, and tourism. They felt the need to do this because this war has highlighted the fact that there are always global energy price fluctuations. The UAE has also attempted to strengthen its connections with big powers like the US and China so that it has backup in case something like this happens to them as well.

Results and Conclusion

COVID-19

Looking at all the data, it is clear that industries associated with travel have faced a significant decline after COVID-19 due to the many different restrictions that were implemented. Also due to all the economic uncertainty that this pandemic caused, many individuals lost their jobs and stopped earning an income. Since a lot of individuals stopped earning, their disposable income had a significant decrease, and they could not afford to travel or have a vacation of any sorts. However, industries that sell necessities and essentials do not have an impact in their sum of volume. This is because these industries sell needs and not wants. Therefore, without these products people cannot live and therefore it will always be required no matter what happens. It is also safe to say that COVID-19 highlighted the vulnerabilities in the UAE economy and now the UAE is working to fix these issues and reinforce it so that if something like this happens again, they are not affected in the same way.

Russia Ukraine War

This war has had both negative and positive impacts on the UAE. On one hand the UAE has had an increase in oil and gas revenue as global prices for energy have increased due to the war. On the other hand, imports have increased in price and therefore increasing costs for businesses. As their costs increase, they pass these costs down to the customers and therefore there is a higher cost of living as goods and services have increased in price. Another effect of this war is that it has heightened regional security concerns for the UAE. This has led the UAE to spend more money on its army and military services. This will help protect the UAE in case a war like this occurs against them. Overall, we can see that global events have had a significant impact on the consumer economy of the UAE.

Works Cited

- 1- Al Mansoori - Alalawi, S., Issa, S.T., Takshe, A.A. and ElBarazi, I. (2022). A review of the environmental implications of the COVID-19 pandemic in the United Arab Emirates. *Environmental Challenges*, 8, p.100561. doi:<https://doi.org/10.1016/j.envc.2022.100561>.
- 2- PricewaterhouseCoopers. "How Resilient Middle East Consumers Have Responded to COVID-19." PwC, May 2020
www.pwc.com/m1/en/publications/how-resilient-middle-east-consumers-have-responded-covid-19.html.
- 3- Shoukry, Ahmed, et al. The War in Ukraine and Inflation Drivers in the GCC: Evidence from the War in Ukraine and Inflation Drivers in the GCC: Evidence from Dubai Muhamad Olimat. 2023.
https://vc.bridgew.edu/cgi/viewcontent.cgi?params=/context/jiws/article/3142/&path_info=10._Rashad__El_Sholkamy____Olimat_The_War_in_Ukraine_and_Inflation_Drivers_in_the_GCC__Evidence_from_Dubai.pdf
- 4- Altamimi, Maryam. "The Impact of Covid-19 Pandemic on Customers's Spending Behavior in the United Arab Emirates." *International Journal of Business and Management*, vol. 17, no. 9, 2023, pp. 1–8, ideas.repec.org/a/ibn/ijbmjn/v17y2023i9p8.html. Accessed 22 Oct. 2024.
- 5- Arab Gulf States Institute in Washington. (Feb 2022). What the Ukraine Crisis Means for Gulf Economies. <https://agsiw.org/what-the-ukraine-crisis-means-for-gulf-economies/>.
- 6- Turak, Natasha. "Dubai Continues to Benefit from Crisis and the Russia-Ukraine War, Damac Chair Says." CNBC, 17 Jan. 2023,
www.cnbc.com/2023/01/17/dubai-benefits-from-crisis-and-the-russia-ukraine-war-damac-chair.html.

AI in the Creative Music Economy: Legal Frameworks and Small Artists By Carly Ding

Keywords

Legal, Generative AI, Creative Economy, Music Industry, Small Artists, Ownership, Freedom, Copyright.

Abstract

The ambiguous nature of generative AI has challenged ownership rights and creative freedom across innovative industries. This paper reviews current literature and summarizes how the current legal framework for AI-generated music impacts the ownership rights and creative freedom of smaller artists, particularly in comparison to larger established artists and labels. Using two case studies, this paper identifies patterns, trends, and potential solutions to address the inequities faced by smaller artists, suggesting changes to empower and improve equity. Findings make it evident that generative AI services and users should focus on clear and transparent terms of service regarding data ownership, while policymakers should enhance the interpretation of copyright law to explicitly address AI-generated content. While the issues raised extend beyond music, with similar cases involving text-based LLMs, the music industry has seen faster litigation, potentially due to the size and power of record labels. This paper focuses on the music sector, recognizing the potential for overlapping case precedents and the need for continued attention across creative industries.

Introduction

The rise of AI-generated music has disrupted the music industry, with smaller artists disproportionately affected in terms of ownership rights and creative freedom. The current laws around AI-generated materials are insufficient to address the unfair treatment of smaller artists, specifically ownership rights. (Ephrat Livni et al.)

The addition of artificial intelligence (AI) has shaped a new era of music production, with generative AI capable of creating music that is increasingly different and unique from man-made compositions (Clair). While this technological advancement holds great potential for the music industry, it has also brought along unprecedented challenges, particularly for original artists—the author of the sampled piece, as opposed to the AI-derivative artist who uses these tools to support composition. Concerns are arising regarding the impact of AI-generated music on ownership rights, copyright protection and creative freedom. Current research and public discussion has primarily focused on the efficiencies, benefits and broader implications of AI's impact (Appel et al.), with limited exploration into copyright and ownership issues. This study aims to delve deeper into aspects in which AI is infringing the rights and freedoms of these independent musicians, and comparing their experiences with those larger more established artists under record labels. Moreover, it will comment upon the role of AI literacy in navigating this evolving field and its potential impact on the future of the music industry.

However, it is important to differentiate between various forms of AI-generated music. Derivative artists utilize AI as a tool to enhance their creative process, often incorporating elements of existing works. AI-assisted derivative artists rely more on AI algorithms to generate musical compositions, though human intervention remains significant in shaping the final output. Solely-AI derivative works, on the other hand, are generated entirely by AI without human input. Understanding these distinctions is essential for evaluating the impact of AI on ownership rights, copyright protection, and creative freedom.

With the issues of generative AI in mind, the research questions of this study are the following:

- How does the current legal framework for AI-generated music impact the ownership rights and creative freedom of smaller artists, particularly in comparison to larger established artists and labels?
- What changes would empower and improve equity for small artists?

The current legal framework surrounding AI-generated music often leans in favor of larger, established artists and major record labels, leaving smaller artists with inadequate protection for their ownership rights and creative freedom. This imbalance creates an unfair playing field, hindering the ability of smaller artists to innovate and thrive in the industry.

Existing laws disproportionately benefit larger, established artists (Palmigiano et al.) in the realm of AI-generated music, leaving smaller artists vulnerable to the exploitation of their creative work (Reed). While recent studies on AI's impact on the creative industries have primarily focused on the broader implications, there is a notable lack of research specifically addressing the concerns of smaller artists. By examining copyright and ownership issues, this research aims to fill this critical gap in the literature.

Methodology

To comprehensively understand the impact of AI on the music industry, particularly for smaller artists, this work will employ a case study approach. A thorough literature review will examine the rise of AI in music, (Deign) exploring its implications for the industry and the specific challenges faced by smaller artists within the context of copyright. Building upon this foundation, in-depth case studies will be conducted to illuminate the practical realities faced by these artists in an AI-driven landscape. This combined approach will provide a robust analysis of the issue, identifying patterns, trends, and potential solutions.

The first section will provide a contextual overview of the rise of AI in music and its impact on the industry. The second section will delve into the legal framework governing AI-generated music, analyzing its strengths and weaknesses in protecting the rights of smaller artists.

Case Study 1: AI-Powered Music Composition Tools and Copyright will focus on the complexities of copyright ownership and attribution in relation to AI-generated music. By

examining the role of human creators and AI algorithms, this case study will shed light on the challenges of determining copyright ownership in AI-generated works.

Case Study 2: AI Companies in Music Making will examine the lawsuit filed by major record labels against AI music startups Udio and Suno, highlighting the growing concerns around copyright infringement in the age of AI-generated music. This case study will explore the legal and business implications of this landmark case, including its potential impact on the future of AI music development.

The paper will conclude by synthesizing the findings from the literature review and case studies to identify key patterns and trends. Based on these insights, the paper will propose recommendations for policymakers, industry stakeholders, and artists to address the identified challenges and create a more equitable landscape for AI-generated music.

Contextual background

The music industry has undergone significant transformation since the early 2000s. Prior to this, the sector was dominated by major labels that controlled artist discovery and promotion. Physical albums were prominent and revenue primarily originated from album sales. Copyright laws were established to protect tangible works in the past, such as physical albums and CDs. However, the rise of digital platforms disrupted this, granting artists unprecedented outreach. Yet, the integration of AI introduces new challenges, such as traditional notions of copyright and ownership (Coote and McCombie).

Legal landscape of AI-generated music refers to the existing laws that regulate the creation (Emancik), promotion and use of music produced using AI. Because generative AI in the field of music is relatively new, there are certain knowledge gaps that we need to identify. AI is affecting the rights and freedom of smaller artists, according to Harvard Business Review (Cremer et al.). That is not the only issue with generative AI in the music industry. It not only impacts smaller artists but also larger, established artists, however, larger artists tend to be less impacted by this, and their legal departments have more say in this. A recent incident where major labels sued AI music startups Udio and Suno, brings forth the complexities of AI's impact on the music industry as a whole (Virginie Berger), (Knibbs). Since the start of AI literacy in the music industries, prejudice has surfaced specifically amongst smaller artists, who fear that AI-generated music could mimic their styles, blurring the lines of originality and ownership. This concern mirrors the anxieties expressed by writers and content creators in the text-based realm, where LLMs like ChatGPT have raised similar questions about copyright infringement and the protection of original works.(Grynbaum and Mac) These parallel developments highlight the broader challenges posed by AI in preserving creative integrity and ensuring fair compensation for creators across different media.

Current issues

The landscape of generative AI has undergone a dramatic transformation since late 2022, with rapid advancements across text, image, and music generation. Initially, the focus was on the

capabilities of large language models (LLMs) like ChatGPT, demonstrating their proficiency in generating human-quality text formats. This rapid progress allowed for a broader exploration of generative AI's potential, extending into image generation with tools like Midjourney (Midjourney).

Specifically, it has now expanded to music generation, with AI models capable of producing increasingly complex compositions. This rapid evolution has introduced an interplay of technological advancement, ethical considerations, and legal challenges (Holloway et al.). Issues such as copyright infringement, artist rights, and the potential for deepfakes have become more common (Hannan) with the rise of generative AI.

Case study 1: AI-Powered Music Composition Tools and Copyright

This case study explores the legal and ethical concerns surrounding AI-powered music composition tools ([Boomy Corporation](#)) such as the challenges of copyright ownership and attribution. By examining the strengths and limitations of these tools, we can have a better understanding for artists, copyright holders, and the music industry as a whole.

In recent years, AI-powered music composition tools have made significant growth, thanks to advancements in machine learning and deep learning techniques. These tools are now capable of generating complex and diverse musical compositions, including melodies, harmonies, and even lyrics; examples of popular AI music generators include AIVA (AIVA), and Boomy (Boomy Corporation). AI-powered music composition tools have advanced significantly, capable of generating original melodies, harmonies, and even lyrics. These tools can adapt to specific styles or moods, creating chord progressions that enhance melodies. Additionally, some can generate lyrics based on given topics or styles. Since the rise of AI in 2022 (OpenAI), the widespread use of large language models has sparked discussions about authorship and ownership. While AI is a powerful tool for generating knowledge and creative ideas, the question of who holds the copyright to AI-generated content remains a complex issue. Studies have shown that the answer often depends on the specific terms and conditions of the AI platform being used. For example, platforms like Boomy, a platform for AI-generated music, has its own terms and conditions (Boomy Corporation, 'Terms of Use') stating that users who input their original work onto their platform and use their AI tools to create new content can retain ownership and copyright of their original work, even if the AI was used to enhance or modify it. However, for other gen AI softwares, they may have different terms and conditions. A contrasting music tool from boomy is: AIVA where they have stated that the copyright holder depends on the plan that you purchase. For example, they have stated that if you're under the free or standard plan, the copyright of the music generated by the platform is owned by the company, but if you get a premium plan the copyright is held by the user (AIVA). While AI-powered compositional tools can boost the creative economy by inspiring human composers to explore new ideas and experiment with different styles, AI could influence changes in legal ownership of generated music, leading to a loss of originality and clarity of ownership in music.

The existing copyright laws are still unclear as AI-generated music is still relatively new to this legal landscape. Traditional copyright principles, such as fair use and originality, were developed for human-created works and may not fully capture the nuances of AI-generated content. In order to determine and protect the work of artists, there is a need to focus on fair use as well as originality of works.

In the field of fair use, AI-generated music for commercial purposes or to create similar works imitating that of other artists made up against fair use of the large music model. The nature of the copyright at work such as whether it be similar melodies, harmonies and lyrics depends on the analysis of its authenticity. A recent incident of an AI-powered company creating a song that mimics that of Drake and The Weeknd has received backlash which can be seen as infringement (Coote and McCombie). In the past, works created for commercial purposes are less likely to qualify for fair use which raises concerns about the use of AI-generated music in commercial settings. As Clifford Chance highlights, "the major concerns for rightsholders (as exemplified by the newly formed Human Artistry Campaign) is that copyright protection for AI-generated music could undercut their catalogs and undermine human creativity" (Coote and McCombie). Here, the argument is that commercially exploiting AI-generated music that potentially mimics existing works could diminish the value of those originals.

Additionally, creating derivative works solely through AI manipulation of copyrighted material might not constitute transformative use, another key factor in fair use analysis. The concept of originality, another cornerstone of copyright protection, presents another layer of complexity. AI-generated music often relies on vast amounts of training data, which may include copyrighted works. This raises questions about whether the resulting AI output can be considered truly original. Harvard Law School Professor Rebecca Tushnet emphasizes that "the originality requirement is particularly tricky when it comes to AI-generated works, because it's not always clear who or what is the author" (Reed). Determining the level of originality in AI-generated music becomes crucial when assessing fair use, as heavily derivative works are less likely to be protected.

To navigate the complexities of AI-generated music, it is essential to explore potential reforms to copyright law. New licensing models and industry standards can play a crucial role in clarifying ownership rights (Firemark) and promoting fair practices: such as the purpose of the use, the nature of the copyrighted work, the amount and substantiality of the portion used, and the effect on the market for the original work. By addressing these areas, policymakers and industry stakeholders can work together to create a legal and ethical framework that supports innovation and protects the rights of all parties involved in the creation and distribution of AI-generated music (Handy).

Case study 2: AI companies in making music

This case studies a lawsuit of major record labels against the AI music startups, Udio and Suno. After generative AI has been introduced in the music industry, copyright infringement has

been on the rise. This case depicts the potential how generative AI can impact and integrate into the music production, not just within but also across the greater music industry.

The advent of AI-powered music generation tools has resulted in a new era in music production, revolutionizing the industry in new ways. But this technological advancement has also given rise to some overwhelming challenges, especially with regard to copyright and ownership. Major record labels suing AI music startups like Udio and Suno is a serious reminder of how fearsomely complicated the legal landscape has become in light of AI merging into the creation of music (Robinson). Central to the case is copyright infringement. The major record labels involved claim that Udio and Suno infringed the copyrights of innumerable songs by using their copyrighted material to train their AI models to create new music (Sato). At issue are basic questions regarding the reach of copyright protection in the digital age and how much AI can draw upon pre-existing works in creating original works. The lawsuit between major record labels and AI music startups Udio and Suno over copyright infringement is ongoing. For their part, Udio and Suno contend that their AI models create original and transformative works that do not infringe on existing copyrights (Nwaokocha). The case is expected to have significant implications for the future of AI-generated music and copyright law. Record labels claim that the use of copyrighted material undercuts the value of their artists' work and may eventually lower demand for original music (Pasteau). This court case may have faced the most controversial issue regarding the use of generative AI in the music industry, firing and offending major record labels such as Universal Music Group, Sony Music Entertainment, Warner Music Group and other music giants (Carras) through “taking an aggressive stance to protect their intellectual property against technology that makes it easy for people to generate music based on existing songs.” (Carman and Metz). While Udio and Suno have demonstrated impressive capabilities in generating high-quality AI-generated music, their actions in allegedly exploiting copyrighted material raise serious concerns. If these companies had operated within the boundaries of copyright law, their AI tools could have been valuable tools for production for musicians and music creators, providing new sparks for creativity and experimentation. Unfortunately, the controversy surrounding their alleged infringement has overshadowed their technological advancements.

In the end, this case has become a very important test of both the challenges and possibilities brought in by AI-generated music. The lawsuit between major record labels and AI music startups Udio and Suno underscores the complex legal and ethical challenges by AI-generated music. It is important now for lawmakers to explore potential innovative licensing models that accommodate the use of AI in music creation. A balance should be struck between innovation and protection of the rights of artists.

Analysis

The two case studies included in the paper have varying perspectives on challenges and opportunities brought forth by AI-generated music. While both case studies are dealing with the

issues of AI and Copyright, they focus on different elements of the larger legal and ethical landscape.

Case study 1 provides an overview of the key legal and ethical implications of AI-generated music. It analyzes the problems of applying traditional copyright principles to works created by AI and gives some possible solutions for overcoming those problems. On the other hand, Case study 2 focuses on the legal fight between major record labels and AI music startups. It illustrates practical difficulties and possible consequences of copyright infringement within AI-created music.

Whereas both case studies have different things to say about challenges and opportunities that AI-generated music creates. Case Study 1 provides more of a broad overview of the legal and ethical issues, whereas Case Study 2 focuses on one single legal issue. Looking at both of these together helps us appreciate the complex issues that are linked to AI and copyright in the music industry.

In short, both case studies illustrate the urgent need under policy for there to be ongoing dialogue and consultation between policy makers, industry producers, and artists regarding the updating of new legislative frameworks around copyright and ethics. With the continued development of AI, serious problems and possibilities about AI-generated music will need to be addressed in a proactive and integrated fashion.

Conclusion

This research has also understood the delicate relation between AI music, copyright issues and especially the rights of the minor artists. Legal analysis, case studies, and literature review have brought out several problems and even some solutions.

Small artists are so much more defenseless against the harmful side effects of the computer generated music, for they do not have weapons or negotiating powers. In this situation, existing copyright provides little comfort and again with the newer *Pioneers of Composition* Devices comes confusion about whose rights over all AI work.

These problems can be solved through revision of the copyright law, development of new licensing schemes and formulation of industry standards. In this course, a more even handed and ecologically sound perspective of the music market could be envisioned. Worked embodiments, leaping of scientific and technical advancements, receptiveness to foreign countries, and raising of moral problems showcase directions for future work.

Acknowledgements

This work was supported by Dr Steve Bullock from the University of Bristol, through the Indigo Research mentorship programme.

Works Cited

- AIVA. 'I Don't Understand the Terms of License'. AIVA Helpdesk, <https://aiva.crisp.help/en/article/i-dont-understand-the-terms-of-license-1wqvh5v/>. Accessed 7 Sept. 2024.
- Appel, Gil, et al. 'Generative AI Has an Intellectual Property Problem'. Harvard Business Review, 7 Apr. 2023. [hbr.org](https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem), <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>.
- Boomy Corporation. 'Homepage'. Boomy, <https://boomy.com/>. Accessed 7 Sept. 2024.
- . 'Terms of Use'. Boomy, <https://boomy.com/terms>. Accessed 7 Sept. 2024.
- Carman, Ashley, and Rachel Metz. Major Record Labels Sue AI Music Generators | TIME. <https://time.com/6991466/record-labels-sue-ai-music-generator-startups/>. Accessed 3 Oct. 2024.
- Carras, Christi. 'Major Record Labels Sue AI Companies for Allegedly Stealing Copyrighted Music'. Los Angeles Times, 24 June 2024, <https://www.latimes.com/entertainment-arts/business/story/2024-06-24/riaa-suno-udio-la-wsuit-ai-copyright-songs-music>.
- Clair, Adam. 'What AI in Music Can — and Can't — Do'. Vox, 5 Aug. 2024, <https://www.vox.com/the-highlight/358201/how-does-ai-music-work-benefits-creativity-production-spotify>.
- Coote, Jonathan, and Don McCombie. 'AI-Generated Music and Copyright'. Clifford Chance, <https://www.cliffordchance.com/content/cliffordchance/insights/resources/blogs/talking-tech/en/articles/2023/04/ai-generated-music-and-copyright.html>. Accessed 7 Sept. 2024.
- Cremer, David De, et al. 'How Generative AI Could Disrupt Creative Work'. Harvard Business Review, 13 Apr. 2023, <https://hbr.org/2023/04/how-generative-ai-could-disrupt-creative-work>.
- Deign, Jason. The Music Industry Is Tuning into AI. <https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2024/m02/the-music-industry-is-tuning-into-ai.html>. Accessed 17 Aug. 2024.
- Ephrat Livni, et al. 'Who Owns the Song You Wrote with AI? An Expert Explains'. World Economic Forum, 18 Aug. 2023, <https://www.weforum.org/agenda/2023/08/intellectual-property-ai-creativity/>.
- Firemark, Gordon. 'Understanding Copyright Law: What You Should Know About Using Artwork Found Online'. Apparelist, 27 June 2023, <https://www.apparelist.com/2023/06/27/understanding-copyright-law-what-you-should-know-about-using-artwork-found-online/>.
- Grynbaum, Michael M., and Ryan Mac. 'The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work'. The New York Times, 27 Dec. 2023, <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>.
- Handy, Jake. 'Navigating AI Music Tools and Copyright: An Overview for Rightsholders'. Pex, 7 May 2024, <https://pex.com/blog/navigating-ai-music-tools-and-copyright-an-overview-for-rightsholders/>.

- Hannan, Nabil. 'The Rise of Deepfakes and What They Mean for Security'. Information Week, <https://www.informationweek.com/machine-learning-ai/the-rise-of-deepfakes-and-what-they-mean-for-security>. Accessed 30 Sept. 2024.
- Holloway, Jim, et al. 'Will Copyright Law Enable or Inhibit Generative AI?' World Economic Forum, 13 Jan. 2024, <https://www.weforum.org/agenda/2024/01/cracking-the-code-generative-ai-and-intellectual-property/>.
- Knibbs, Kate. US Record Labels Sue AI Music Generators Suno and Udio for Copyright Infringement. <https://www.wired.com/story/ai-music-generators-suno-and-udio-sued-for-copyright-infringement/>. Accessed 14 Aug. 2024.
- Midjourney. Home Page. <https://www.midjourney.com/>. Accessed 30 Sept. 2024.
- Nwaokocha, Amaka. 'Suno Defends AI Training with Copyrighted Music amid RIAA Lawsuit'. Cointelegraph, 2 Aug. 2024, <https://cointelegraph.com/news/suno-ai-training-fair-use-riaa-lawsuit>.
- OpenAI. Introducing ChatGPT. <https://openai.com/index/chatgpt/>. Accessed 7 Sept. 2024.
- Palmigiano, Paolo, et al. 'Artistic Control or Market Manipulation? Enforceability of Non-Resale Clauses'. TaylorWessing, 18 June 2024, <https://www.taylorwessing.com/en/insights-and-events/insights/2024/06/artistic-control-or-market-manipulation>.
- Pasteau, Cyril. 'AI Is Shaking up the World of Music Production - Hello Future Orange'. Hello Future, 18 June 2024, <https://hellofuture.orange.com/en/udio-suno-ai-generated-music-is-already-competing-for-low-added-value-orders/>.
- Reed, Rachel. 'AI Created a Song Mimicking the Work of Drake and The Weeknd. What Does That Mean for Copyright Law?' Harvard Law School, <https://hls.harvard.edu/today/ai-created-a-song-mimicking-the-work-of-drake-and-the-weeknd-what-does-that-mean-for-copyright-law/>. Accessed 5 Sept. 2024.
- Robinson, Rob. 'Legal Battles Erupt as Music Giants Sue AI Firms Over Copyright Infringement'. ComplexDiscovery, 1 July 2024, <https://complexdiscovery.com/legal-battles-erupt-as-music-giants-sue-ai-firms-over-copyright-infringement/>.
- Sato, Mia. Major Record Labels Sue AI Company behind 'BBL Drizzy' - The Verge. <https://www.theverge.com/2024/6/24/24184710/riaa-ai-lawsuit-suno-udio-copyright-umg-sony-warner>. Accessed 23 Sept. 2024.
- Semancik, Alex. 'How AI Is Transforming the Creative Economy and Music Industry'. OHIO News, 4 Mar. 2024, <https://news.ohio.edu/news/2024/04/how-ai-transforming-creative-economy-and-music-industry>.
- Virginie Berger. Why Major Labels Are Suing AI Music Startups Udio And Suno For Mass Copyright Infringement. <https://www.forbes.com/sites/virginieberger/2024/06/27/why-major-labels-are-suing-ai-music-startups-udio-and-suno-for-mass-copyright-infringement/>. Accessed 14 Aug. 2024.

The Future of Huntington's Disease Treatments Through Genetic Editing

By Adithi Chakkera

Abstract

This is a research article discussing how various genetic editing strategies can be used as a treatment for Huntington's disease. The topics covered in this paper are RNA interference, CRISPR-Cas9, ASOs, and TALEN editing techniques. This article discusses advantages and disadvantages, mechanism of action, and clinical trials and development stages of each of these strategies in order to discuss their potential as a treatment.

Introduction

Huntington's disease is a neurodegenerative disease that results in the damage of brain cells over time. This results in loss of function of brain cells, and can lead to a decline in movement and cognition (8). Huntington's disease is an inherited, autosomal dominant disorder caused by a mutation in the huntingtin gene, which causes repeats in CAG DNA segments, containing cytosine, adenine, and guanine (8). Individuals with 36 or more such repeats are likely to develop the disease. This mutation occurs on chromosome 4, producing a mutant protein that causes neural damage on parts of the basal ganglia. This can result in transcriptional dysregulation, impaired protein degradation, mitochondrial dysfunction, synaptic dysfunction, and neuron death (10). The disease gets progressively worse over time, with the worst symptoms emerging during mid adulthood between the ages of 30 to 50. After a certain point, the disease becomes terminal, with a life expectancy of 10-25 years after symptoms develop (11). This emphasizes the need for more effective treatments for the disease. About 30,000 people in the U.S have the disease, with another 200,000 individuals at risk for developing it.

As Huntington's disease is a genetic disease, scientists are currently exploring the possibility of genetic editing as a mechanism of treating the disorder. This may be a potential solution due to the ability to silence the mutation on the gene encoding for the huntingtin protein, responsible for causing the disorder (3). Many forms of genetic editing are being explored as a treatment for Huntington's disease, and each has a different route to effectiveness on treatment (5). The most common forms of gene editing being explored as potential treatments are Transcription Activator-Like Effector Nuclease (TALEN), Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR-Cas9), RNA Interference, and Antisense Oligonucleotide (ASO).

Gene editing techniques as treatments for Huntington disease

1. RNA interference

The buildup of the mutant protein as a result of CAG repeats disrupts many bodily functions, and the most dangerous impact is neurodegeneration, as seen in Huntington's disease. RNA interference is a type of treatment for Huntington's disease involving the targeting of the mutant gene causing the disease by small interfering RNAs (13). The mutant gene is degraded

and becomes destroyed by the siRNAs. This is because the siRNAs will reduce the amount of abnormal huntingtin protein in the cells. As the siRNAs target the gene, the mRNA of the HTT gene is degraded, stopping transcription of the mutant HTT proteins that cause neural damage (4). RNAi lowers HTT levels in the brain by reducing the HTT protein's activity, leading to improved motor function, lifespan, and survival of neurons (1). RNAi treatments are currently in stages 1 and 2 of development with certain drugs being developed to administer the siRNAs into the body. Although its effectiveness for Huntington's is still being developed, there have already been approved treatments using RNAi for other genetic disorders (14). It is a promising treatment as it has shown to be effective on its test subjects. The main delivery mechanism for the siRNAs is via IV injection (14). However, this comes with many challenges, with the most notable being the blood brain barrier. Scientists are currently still working on ways to penetrate this barrier in order to effectively deliver drugs to the brain. The effectiveness of this treatment is being tested by administering the siRNAs to cultured mammalian cells (1). What also makes this an appealing form of treatment is that it is relatively cost efficient compared to other treatments such as CRISPR-Cas9. There are certain risks posed with RNA interference. Off-target effects are likely, when the drug binds to molecules in the body that it was not supposed to bind to. This can cause genetic mutations and unintended gene silencing, resulting in permanent alterations to the genome.

2. CRISPR-Cas9

CRISPR-Cas9 works through a guide RNA, which directs the Cas9 enzyme to the target stand of DNA. Once targeted, the Cas9 enzyme will cut the DNA and allow for genetic editing and modifications to be made. Scientists typically repair or inactivate the gene, which results in reduced or complete haltage of the genetic expression (12). This has revolutionized Huntington's disease treatments because of the strategy's ability to inactivate the mutant HTT protein (2). When using CRISPR, scientists are likely to use allele specific editing, which allows the treatment to only target the mutant allele of the gene while preserving the wild type (12). Like RNA interference, CRISPR treatments also pose the risk of off-target effects. Therefore, the development of specificity and precision of the guide RNA is especially important in order to prevent such errors. Some early-stage clinical trials are exploring CRISPR-Cas9 as a therapy for Huntington's, though it's still in experimental phases for this specific disease (12). Researchers have used mouse models that mimic the genetic mutation responsible for Huntington's to test the efficacy of CRISPR-Cas9 in targeting and editing the faulty *HTT* gene. In several studies, scientists injected CRISPR-Cas9 components directly into the brain or used viral vectors to deliver the gene-editing tool. Potential benefits include long-term effects from a single treatment and the ability to directly address the root genetic cause of the disease (2). However, challenges include ensuring precise targeting to avoid off-target effects, possible immune reactions to the treatment, and the difficulty of delivering CRISPR-Cas9 into brain cells (12). While CRISPR-Cas9 shows significant promise, further research and clinical testing are essential to ensure its safety and efficacy for Huntington's disease.

Talen

TALENs (Transcription Activator-Like Effector Nucleases) work by using proteins to cut DNA (7). The TALEN is a fusion of two proteins, a DNA binding domain which will target the mutant gene and the DNA cleavage domain, which will cut the gene once targeted, working in a similar way to CRISPR (16). However, TALEN is more based on interactions occurring between DNA molecules whereas CRISPR is more RNA based. TALEN is being explored as a treatment for Huntington's because of one of its main benefits, where the TALENs have a great amount of specificity through their functioning. Their specificity, as well as their flexibility in selecting and targeting a gene makes them a more ideal choice as a treatment. This is crucial for any form of genetic editing treatment. Specifically for Huntington's disease, studies conducted in mice have shown that TALENs have been able to significantly reduce HTT levels, suggesting their efficiency as a treatment (16). However, TALENs have challenges like every other treatment. The main challenge associated with TALENs is efficient delivery. TALEN proteins are large in size and complex in structure, which make it difficult for them to penetrate the blood brain barrier and efficiently target genes (7). Despite the challenges, the ability of TALENs to selectively target genes with great accuracy makes it a promising treatment for Huntington's disease that is further being explored and developed by research scientists.

ASO

Antisense Oligonucleotide, or ASOs are small pieces of DNA that will bind to the mutant HTT protein and reduce its production (8). By reducing the amount of the mutant HTT protein in the body, ASOs help reduce Huntington's symptoms. The DNA strands of ASO work by binding to the mRNA coding strand for the mutant protein. Once bound to the mRNA, translation of the protein is blocked (15). One of the most known clinical trials performed for this treatment is done with tominersen (9). Tominersen is an injectable drug that is currently in development, and is investigated as a Huntington's disease treatment drug meant to slow down disease progression and increase life expectancy of patients (9). The drug injects ASOs into the body, which will help block the translation of the HTT protein. However, a recent study showed that the drug reduced both normal and mutant protein levels, which was not the intended outcome (6). Therefore, scientists are still working to develop the drug and properly deliver ASOs into the body (6). ASOs have become a promising treatment for Huntington's disease because of many factors, including their longer half-lives, selective targeting, and their wide distribution across the nervous system (15). Therefore, it is important that scientists work to counter the risks and disadvantages of this treatment, including their difficulty in mechanism of delivery, off target effects, and potentially disturbing side effects (6).

Conclusion

Genetic editing strategies such as CRISPR-Cas9, ASOs, TALENs, and RNA interference are some of many genetic editing strategies that have revolutionized treatment for neurodegeneration as seen in Huntington's disease. While many of these genetic editing

strategies are currently still in development, clinical trials have shown that genetic editing is a promising treatment for many Huntington's disease patients' ability to silence mutant proteins, and reduce its expression and therefore disease symptoms.

Works Cited

- Aguiar, S., van der Gaag, B., & Cortese, F. A. B. (2017). RNAi mechanisms in Huntington's disease therapy: siRNA versus shRNA. *Translational Neurodegeneration*, 6(1).
<https://doi.org/10.1186/s40035-017-0101-9>
- Biswas, T. (2023). Synthego | Full Stack Genome Engineering. www.synthego.com.
<https://www.synthego.com/blog/crispr-huntington-research>
- Duan, W., Urani, E., & Mattson, M. P. (2023). The potential of gene editing for Huntington's disease. *Trends in Neurosciences*, 0(0). <https://doi.org/10.1016/j.tins.2023.02.005>
- Gao, K., & Huang, L. (2008). Nonviral Methods for siRNA Delivery. *Molecular Pharmaceutics*, 6(3), 651–658. <https://doi.org/10.1021/mp800134q>
- Gene Editing Therapy in Early Stages of Huntington's Disease May Slow Down Symptom Progression, Mouse Study Shows. (n.d.). www.hopkinsmedicine.org.
<https://www.hopkinsmedicine.org/news/newsroom/news-releases/2021/07/gene-editing-therapy-in-early-stages-of-huntingtons-disease-may-slow-down-symptom-progression-mouse-study-shows>
- Lauffer, M. C., van Roon-Mom, W., & Aartsma-Rus, A. (2024). Possibilities and limitations of antisense oligonucleotide therapies for the treatment of monogenic disorders. *Communications Medicine*, 4(1), 1–11. <https://doi.org/10.1038/s43856-023-00419-1>
- Malankhanova, T. B., Malakhova, A. A., Medvedev, S. P., & Zakian, S. M. (2017). Modern Genome Editing Technologies in Huntington's Disease Research. *Journal of Huntington's Disease*, 6(1), 19–31. <https://doi.org/10.3233/jhd-160222>
- Mayo Clinic. (2020, April 14). Huntington's disease . Mayo Clinic; Mayo Clinic.
<https://www.mayoclinic.org/diseases-conditions/huntingtons-disease/symptoms-causes/syc-20356117>
- McColgan, P., Alpa Thobhani, Boak, L., Schobel, S., Nicotra, A., Palermo, G., Trundell, D., Zhou, J., Schlegel, V., Patricia Sanwald Ducray, Hawellek, D. J., Dorn, J. F., Cédric Simillion, Lindemann, M., Wheelock, V., Dürr, A., Anderson, K. E., Long, J. D., Wild, E. J., & G. Bernhard Landwehrmeyer. (2023). Tominersen in Adults with Manifest Huntington's Disease. *The New England Journal of Medicine*, 389(23), 2203–2205.
<https://doi.org/10.1056/nejmc2300400>
- National Institute of Neurological Disorders and Stroke. (2022). Huntington's Disease. National Institute of Neurological Disorders and Stroke.
<https://www.ninds.nih.gov/health-information/disorders/huntingtons-disease>
- NORD. (2015). Huntington's Disease - NORD (National Organization for Rare Disorders). NORD (National Organization for Rare Disorders); NORD.
<https://rarediseases.org/rare-diseases/huntingtons-disease/>
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), 2281–2308.
<https://www.nature.com/articles/nprot.2013.143>

- Silencing Huntington's Disease – RNA Therapeutics Institute – UMMS. (2013, November 3). UMass Chan Medical School. <https://www.umassmed.edu/rti/therapeutics/huntingtons/>
- Tiemann, K., & Rossi, J. J. (2009). RNAi-based therapeutics—current status, challenges and prospects. *EMBO Molecular Medicine*, 1(3), 142–151. <https://doi.org/10.1002/emmm.200900023>
- Tuszynier, A. (2024, January 17). Antisense therapy fails to delay the progression of Huntington's disease - Mabion. Mabion. <https://www.mabion.eu/science-hub/science-news/antisense-therapy-fails-to-delay-the-progression-of-huntingtons-disease/>
- What You Need to Know About TALENs. (n.d.). Science Speaks. <https://www.isaaa.org/blog/entry/default.asp?BlogDate=3/24/2021>

How can Neurocircuitry Help to Improve the Understanding and Treatment of Depression?

By Alexis Lavelly

Abstract

Depression is a mood disorder characterized by low mood, lack of interest or pleasure, irritability, sleep difficulties, appetite fluctuations, and suicidality, all of which can severely impact functioning. This review explores the neural circuitry of depression, aiming to enhance our understanding and treatment of the disorder. Neuroimaging studies reveal significant structural and functional differences between individuals with Major Depressive Disorder (MDD) and healthy controls, including reduced gray matter volume—particularly in the hippocampus and prefrontal cortex—and abnormalities in white matter integrity. Functional neuroimaging indicates reduced activation in the anterior cingulate cortex (ACC) and impaired connectivity within critical networks like the default mode and salience networks. While chronic and episodic depression share some neural circuitry, they also exhibit distinct profiles; chronic depression is often linked to reduced functional connectivity and variations in memory performance. Psychological therapies such as Cognitive-Behavioral Therapy (CBT) and Mindfulness-Based Cognitive Therapy (MBCT) play a significant role in reshaping the neural circuits involved in emotional regulation. In contrast, antidepressant medications, especially selective serotonin reuptake inhibitors (SSRIs), contribute to recovery by decreasing amygdala activation in response to negative stimuli and improving connectivity among key brain regions. Additionally, some individuals experience natural recovery without formal treatment, but research on this phenomenon is limited. Understanding it is crucial for uncovering resilience and adaptive mechanisms that could enhance therapy and improve outcomes. This review highlights the need for further research into the shared and unique neurobiological mechanisms of different forms of depression, as well as the psychosocial factors influencing recovery.

Keywords depression, structural neuroimaging, fMRI, persistent depressive disorder

How Can Neurocircuitry Help to Improve the Understanding and Treatment of Depression?

Depression, known as major depressive disorder or clinical depression, is a mood disorder characterized by low mood, and known for producing feelings of persistent doubt or sadness. Symptoms of the disorder include consistently depressed mood, difficulty sleeping, abnormal fluctuations in appetite, feelings of worthlessness, lack of energy, poor concentration, increased feelings of irritability, suicidal thoughts or behaviors, and anhedonia (otherwise known as lack of pleasure). To be diagnosed with MDD, the *Diagnostic and Statistical Manual of Mental Disorders Fifth Edition* (DSM-5) states an individual must experience 5 or more symptoms, including either depressed mood or anhedonia. Depression can affect all people, and is typically characterized by symptoms lasting 2 or more weeks. Severe symptoms of depression can have serious implications on people's ability to function, including difficulty navigating daily tasks and impaired functioning in work, school, and social relationships.

Episodic vs. Chronic Depression (MDD and PDD)

Major Depressive Disorder (MDD) and Persistent Depressive Disorder (PDD), otherwise known as episodic and chronic depression respectively, are both forms of depression separated by a few key differences. Episodic depression (MDD) typically lasts for a duration of at least two weeks. These episodes are typically separated by a period of at least two months, and follow the symptoms of Major Depressive Disorder. However, PDD typically involves less intense symptoms, with its severity being described as mild to moderate. According to the DSM-5, symptoms of Persistent Depressive Disorder include abnormal appetite, insomnia or hypersomnia, low self esteem, chronic fatigue, hopelessness, or poor concentration and decision making. While not as extreme as MDD, PDD symptoms last for at least 2 years, completely taking over and transforming the patient's lifestyle. To be diagnosed with PDD, at least 2 symptoms must be present for 2 or more years. Overall, PDD symptoms are continuous, whereas MDD consists of shorter and more severe episodes broken up by periods of recovery.

Outcomes and Prevalence of Depression

As a disorder, depression has the ability to significantly decrease quality of life, and impair a person's ability to navigate daily stressors (Brody, 2021). In some cases, persistent depression can even lead to physical health issues, increasing the chances of chronic pain, inflammation, or even heart attack or stroke (Kang et al., 2017; Zunszain et al., 2013). Untreated depression can rapidly intensify over time, leading to more severe symptoms and disrupting personal relationships, health, and a person's productivity and ability to work. When left undiagnosed, people can be more likely to engage in risky or harmful behavior, including alcohol or substance abuse, self destructive behaviors, or suicidal tendencies (Brody, 2021).

Depression is one of the most prevalent mental health conditions, affecting nearly 21% of adults in the United States, with those rates rising in young adults between the ages of 18 and 29 (Hastin et al., 2018; Villarroel et al., 2020). According to the National Institution of Mental Health, MDD is among the leading causes for disability in the U.S, affecting more than 21 million adults every year. However, despite the vast presence, only about 34% of adolescents and 40 to 57% of adults received disorder specific treatment for MDD, with those rates fluctuating depending on gender or ethnicity (Avenevoli et al., 2014).

What is Neuroimaging?

Within recent decades, neuroimaging has grown to be a critical tool both in understanding brain function and studying neurological disorders (Yen et al., 2023). Neuroimaging is a field of neuroscience used to understand both the function and structure of the nervous system. Within neuroimaging, neurocircuitry refers to structural and adaptive communication between separate neural regions in the brain, each acting on specific functions when activated (Thibaut, 2018). Functional magnetic resonance imaging (fMRI), and electroencephalography (EEG) are both noninvasive techniques used to observe changes in brain activity (Yen et al., 2023). With the assistance of neuroimaging, researchers gain the ability to

better understand biological mechanisms of depression, as well as the behavior of large scale brain networks (Thibaut, 2018; Dai et al., 2019). When clinically applied, mental disorders could be diagnosed more efficiently and accurately, and individual response to treatment can be better predicted through the visualization of potential biomarkers (Kottaram et al., 2018; Masdeau, 2011). In this review paper, I explore the primary research question: How can neurocircuitry help to improve the understanding and treatment of depression?

Literature Review

Anatomical Differences in Brains Affected by Depression

Within neuroimaging, structural MRI's are utilized to study both gray and white matter in the brain, providing qualitative and quantitative data to describe shape, size, and soundness of various anatomical regions (UC San Diego School of Medicine, Center for Functional MRI, 2024). With the progression of MRI technology, researchers are able to view high-resolution images of brain structure, therefore gaining the ability to both visualize activation patterns and gather region-specific functional signal data (Zeng et al., 2022; UC San Diego School of Medicine, Center for Functional MRI, 2024). When analyzing structural MRI's, the nervous system is divided into two distinct regions: gray matter, which serves as the control center and consists of neuron cell bodies, and white matter, which consists of axons coated in myelin sheaths that transmit essential information between different brain regions (Rogers et al., 2016; Zatorre et al., 2012). Gray matter is crucial for processing information in the brain, as the nerve cell bodies it contains are responsible for not only sensory perception, but also muscle control, and higher cognitive functions such as decision-making, memory, and self-control (Rogers et al., 2016; Zatorre et al., 2012). The main function of white matter, in contrast, is to ensure the efficient transfer of information from various regions of gray matter within the central nervous system (Fields, 2008; Reddick et al., 2006). A decrease in white matter volume and connectivity can disrupt the transmission of information between brain regions, resulting in deficits in attention, memory, executive function, and general cognitive abilities (Fields, 2008; Reddick et al., 2006). To study gray matter structures, voxel-based morphometry (VBM) assesses the volume, surface area, thickness and shape of gray matter, including the subcortical structures and the cerebral neocortex (Dai et al., 2019; UC San Diego School of Medicine, Center for Functional MRI, 2024). VBM divides the brain into gray matter, white matter, and cerebrospinal fluid, and transforms these specific images into a standardized template space, where a smoothing process is then applied (Nemoto, 2017). When examining white matter, macrostructural white matter integrity is evaluated through measuring the volumes of normal and abnormal matter, helping to identify signs of inflammation, edema, or demyelination (UC San Diego School of Medicine, Center for Functional MRI, 2024). To measure the integrity of matter, a type of MRI called diffuser tensor resonance imaging, or DTI, is used to view and map white matter tracts in the brain, enabling researchers to examine networks throughout brain regions, as well as their structures and functions (Boucher et al., 2020; Yen et al., 2023). To accurately determine the

microstructural changes in white matter of patients compared to healthy controls, Tract-Based spatial statistics (TBSS) is used to analyze DTI data (Dai et al., 2019). TBSS focuses on measuring fractional anisotropy (FA), depicting the degree of anisotropy, and acting as a biomarker of white matter integrity (Vandermosten et al., 2012). This macrostructural data complements microstructural diffusion weighted MRI, offering a thorough view of white matter health (UC San Diego School of Medicine, Center for Functional MRI, 2024).

Hippocampal volume reduction or atrophy is the most common structural fluctuation in patients with depression, typically becoming more severe in the cornu ammonis, dentate gyrus, and subiculum (Arnone et al., 2012; Cole et al., 2011; Roddy et al., 2019). Within the nervous system, the hippocampus plays an integral role in both short and long term memory, emotion processing, and learning. Although it's unclear whether hippocampal atrophy is a direct result of depression, or simply a correlation, results show that they mutually worsen with time (Buddeke et al., 2017; Den Heijer et al., 2011). Additionally, the cingulate cortex, a member of the limbic system responsible for various cognitive and emotional processes, appears smaller in depressed patients when compared to typical controls (Rodriguez-Cano et al., 2014; Wise et al., 2017). Furthermore, when examining the frontal lobe, studies indicate the particular significance of medial prefrontal cortex, frontal cortex, and dorsolateral prefrontal cortex atrophy, as those regions are heavily important in emotional regulation and memory (Bludau et al., 2016; Van Eijndhoven et al., 2013; Zhao et al., 2014; Bludau et al., 2014). Similarly to atrophy in the frontal lobe, when compared to healthy controls, the bilateral putamen and left thalamus in depressed patients remains smaller, leading to difficulty processing memory, managing emotions, and transmitting information across brain regions (Lu et al., 2016; Zhang et al., 2016).

In depressed patients, abnormal white matter occurs frequently, potentially due to impaired cerebral blood flow and regulation (CBF), blood-brain barrier operation (BBB), inflammation, genes, and external factors (Liao et al., 2013; He et al., 2022). However, when treated or medicated, white matter FA values in patients with depression can vary due to clinical variables including symptom severity and clinical assessment (He et al., 2022). When compared to a healthy control group, the FA values of white matter in depressed patients were significantly lower in the right cerebellar hemisphere and bilateral superior longitudinal plasma, with values plummeting abnormally low in the corpus callosum: a structure consisting of white matter tracts that connects the left and right central hemispheres (Jiang et al., 2017; Han et al., 2014). White matter research suggests that increased severity of symptoms in depressed patients is associated with white matter hyperintensity, indicating a greater likelihood for microstructural alterations (Cole et al., 2012).

Functional Differences in Brains Affected by Depression

When studying how Major Depressive Disorder (MDD) functions, researchers most commonly use functional magnetic resonance imaging (fMRI), a method that indirectly assesses brain activity and function (Glover, 2011; Pilmeyer et al., 2022). This technique is primarily used in both resting state and task-based studies, and is able to provide valuable insight into

underlying abnormal brain functions associated with psychiatric disorders (Glover, 2011; Pilmeyer et al., 2022). The fMRI is greatly based on MRI techniques, but additionally uses blood oxygenation level-dependant (BOLD) contrast to measure variations in oxygenation concentration within the brain, therefore allowing researchers to view abnormal changes in brain activity during both resting and task-based conditions (Dai et al., 2019; Glover, 2011). Within the field of neuroscience, studies on MDD have utilized fMRI's to identify biomarkers, leading to increased understanding of the functionality and connectivity of brain regions and networks in depressed patients when compared to healthy controls (Dunlop et al., 2017).

Resting-state fMRI (rs-fMRI) is a method used to measure inherent brain activity during periods of rest, providing insight into the brain's natural function, as well as the interaction between functional networks and cognitive processes (Raimondo et al., 2021). Additionally, through removing external stimuli, rs-fMRI's focus on brain activity during rest and idle moments, such as mind-wandering or rumination, particularly highlighting the default mode network's connection to these introspective processes (Rosenbaum et al., 2017). During a resting-state scan, participants are typically asked to simply lie in the scanner either with their eyes closed or focused on a fixation point displayed on a screen. When examining resting state activity in depressed patients, studies have found less activity in the anterior cingulate cortex (ACC) and the right inferior frontal gyrus compared to healthy controls (HCs), which is linked with MDD neuropathy (Brakowski et al., 2017). When observing resting-state, scientists measure functional connectivity— a value measuring the correlation between activity in two or more regions within the brain, and therefore suggesting the collaboration of multiple networks to support various psychological functions (). In MDD, changes in brain connectivity predominantly involve the default mode network (DMN), the salience network (SN), and the central executive network (CEN) – all major brain regions that play critical roles in processing stimuli and managing cognitive control (Brakowski et al., 2017). The DMN, which consists of the dorsal medial prefrontal cortex, posterior cingulate cortex, precuneus and angular gyrus, becomes active when a person is unfocused or not actively engaged in any specific task, while the SN is responsible for filtering stimuli, emotion regulation, behavior control and managing self-awareness. Finally, the CEN, a dominant brain network, directs attention, manages the processing of external stimuli, processes memory information, and is a key component in decision making. Additionally, enhanced dynamic resting state functional connectivity (RSFC) between the medial prefrontal cortex (MPFC) and the insula has been associated with an increase in recent rumination, suggesting that these atypical patterns of fluctuating communication between brain regions may directly lead to symptoms of depression (Brakowski et al., 2017).

In task-based fMRI studies of depression, participants engage in tasks designed to activate brain regions or circuits that are thought to be impaired in MDD (Li et al., 2022). These tasks are designed to assess various aspects of emotional and cognitive functions and can be classified into three main categories. The first category includes emotional processing tasks, such as the angry face processing task and the emotional regulation task. The second category encompasses reward learning and valuation tasks, including the reward guessing task, monetary

reward task, probabilistic selection task, and probabilistic reward task, among others. The third category involves cognitive tasks, such as working memory tasks and the Tower of London paradigm (Li et al., 2022).

The amygdala, a brain region frequently associated with the pathophysiology of mood disorders such as MDD (Whalen et al., 2002), is known to respond to both positive and negative emotional stimuli (Sergeje et al., 2008), and plays a key role in modulating the perceived significance of stimuli and events (Davis and Whalen, 2001). Research has demonstrated that activity in limbic regions, such as the ACC and amygdala, as well as in frontal cortical areas like the dorsolateral prefrontal cortex (dlPFC), is disrupted during the viewing or processing of emotional stimuli, particularly facial expressions, in individuals with depression (Stuhrmann et al., 2011; Ebneabbas et al., 2021). Additionally, decreased functional connectivity between the amygdala and the ventrolateral prefrontal cortex (vlPFC) has been reported in patients with MDD (Hacimusalar et al., 2018; Lener et al., 2015). Furthermore, abnormal increases in activation have been observed in several brain regions, including the right ventrolateral prefrontal cortex (vlPFC), orbitofrontal cortex (OFC), bilateral middle temporal gyri (MTG), superior temporal gyri (STG), and medial frontal gyrus (MFG), as well as the amygdala, striatum, parahippocampal gyrus, cerebellum, fusiform gyrus, and ACC (Li et al., 2022). Conversely, decreased activation was noted in the dorsal anterior cingulate cortex (dACC), striatum, nucleus accumbens (NAcc), caudate, putamen, and globus pallidus (Li et al., 2022), demonstrating that while the activity of limbic regions may increase, brain activity commonly decreases in frontal regions responsible for cognition processes (Pilmeyer et al., 2022). In cognitive tasks, depressed individuals frequently exhibit increased prefrontal activity, even though their performance scores are comparable to those of healthy controls (HCs) (Ebmeier et al., 2006; Thomas et al., 2009). In contrast, reduced activity in these regions is commonly observed alongside poor task performance (Ebmeier et al., 2006; Thomas et al., 2009). These findings suggest that depression is linked to cortical inefficiency, which is partially offset by heightened brain activation in these regions (Pilmeyer et al., 2022).

Neurobiological Differences in Chronic vs Episodic Depression

Research has explored potential differences between chronic and episodic depression through comparing individuals with persistent depressive disorder to those experiencing a first depressive episode. Persistent Depression (PDD) and First Depressive Episode (FDE) are both forms of Major Depressive Disorder (MDD), associated with suicidal ideation, and physical or depressive symptoms (Roca et al., 2011; Lin et al., 2021; Zu et al., 2021). However, while studies have found no significant difference in anxiety levels between PDD and FDE groups, variations in memory performance have been observed, and possibly linked to reduced executive functioning in PDD (Zanatta et al., 2019).

In the DSM-5-TR, which represents the latest American classification system for mental health disorders as of 2022, the previously distinct diagnoses of "dysthymia" and "chronic major depression" have been consolidated into a single category labeled "persistent depressive

disorder" (Bschor et al., 2014). In the ICD-10 classification, "dysthymia" refers to a long-standing depressive syndrome that is typically less severe than what is defined as major depression, and includes chronic forms like PDD (Bschor et al., 2014). Forms of chronic depression such as PDD might be associated with low serotonin levels and are commonly triggered by traumatic events or a series of negative experiences (Rodrigues, 2022). Additionally, many individuals with PDD also experience at least one episode of major depression during their lifetime— a phenomenon often referred to as "double depression" (Rodrigues, 2022).

Neuroimaging studies have demonstrated that individuals with dysthymia exhibit notably reduced activation in the dorsolateral prefrontal cortex compared to healthy controls (Ravindran et al., 2009). Conversely, these individuals display increased activation in the amygdala, ACC, and insula, with these differences becoming more pronounced when processing negative stimuli compared to positive (Ravindran et al., 2009). Such altered neural activation patterns might indicate a common underlying mechanism for depressive disorders overall, while some specific changes (outlined below) may relate to the unique symptoms and chronic nature of dysthymia (Ravindran et al., 2009). These findings are significant, considering that dysthymia, despite its seemingly mild presentation, is still often viewed as character pathology rather than a serious mental health condition (Ravindran et al., 2009).

Research of the structural brain differences between chronic and episodic depression have uncovered distinct variations in cortical thickness, hippocampal volume, and neurotrophic factors, shedding light on the diverse ways these forms of depression impact brain anatomy. In a study examining unmedicated patients within a first episode, researchers found increased cortical thickness in several brain regions, including the temporal pole, the caudal anterior and posterior cingulate cortex, and the rostral dorsolateral prefrontal cortex (DLPFC) compared to individuals without MDD (Qui et al., 2014). However, reduced cortical thickness in the medial orbitofrontal gyrus was also observed, indicating possible variation in dendritic branching or myelination processes near gray-white matter boundaries (Sowell et al., 2004). While alterations in cortical volume are primarily influenced by cortical surface area rather than thickness, cortical thickness measurements still reveal important details about disease-related changes in brain structure, including neuronal density, organization, and support (Qui et al., 2014). While variations in cortical thickness and neural circuitry have been observed, no significant overall differences in hippocampal gray matter volume were found between patients with FEMDD and healthy controls (Frodl et al., 2002). However, post hoc analysis detected notable left-right asymmetry and a reduction in hippocampal white matter volume when compared to healthy controls, possibly resulting from acute factors such as vascular hypoperfusion, or chronic factors like neuronal cell loss or atrophy (Frodl et al., 2002). Furthermore, lower levels of brain-derived neurotrophic factors (BDNF) in response to stress could contribute to hippocampal atrophy and lead to the degeneration of stress-sensitive CA3 pyramidal neurons in more severe cases (Frodl et al., 2002).

To assess brain activity during a patient's first episode of MDD (FEMDD), resting-state functional MRI's (rs-fMRI) are often used, offering valuable insight into the early stages of the

disorder (Zanatta et al., 2019). Using this technology, along with diffusion tensor imaging (DTI), researchers have identified changes in neural circuitry and connectivity during a first episode, revealing alterations in brain network connectivity and structural pathways associated with the disorder's initial onset (Zhang et al., 2011; Tao et al., 2013; Zhu et al., 2011). These irregularities are located in regions supporting the cortical-limbic dysregulation model of MDD, with notable differences in functional connectivity (FC) ranging across the DMN, CCN, and AN in various brain regions including the frontal, parietal, temporal, and precuneus lobes, as well as the cerebellum (Zanatta et al., 2019). Abnormalities observed in FEMDD patients were consistent with those found in individuals with persistent depression, suggesting the presence of similar dysfunctions early in the disorder's formation (Zanatta et al., 2019). The recurrent depressive episode (RDE) group generally exhibited reduced functional connectivity (FC) within the default mode network (DMN), cognitive control network (CCN), and affective network (AN), compared to the first depressive episode (FDE) group (Zanatta et al., 2019). In comparison to healthy controls (HC), the FDE group also displayed decreased FC in these networks, while the RDE group depicted reduced FC only within the DMN and AN (Zanatta et al., 2019). Notably, within the DMN—associated with mind wandering and rumination—the RDE group had increased FC between the left and right posterior cingulate cortices (PCC) and the right inferior temporal gyrus, unlike the FDE group, which displayed impaired FC in these regions (Zanatta et al., 2019).

In the Cognitive Control Network (CCN), which is integral for high-level cognitive functions such as executive control, attention, working memory, and decision-making, the RDE group exhibited significantly enhanced functional connectivity between the left dorsolateral prefrontal cortex (DLPFC) and the right precuneus, when compared to first episode depression (Zanatta et al., 2019). Additionally, the RDE group displayed increased connectivity between the right DLPFC and both the left precentral gyrus and the left posterior cerebellar lobes, relative to the FDE group (Zanatta et al., 2019). Abnormal FC between the CCN and post-DMN may reveal distinct neuropathological mechanisms underlying RDE compared to FDE (Zanatta et al., 2019).

Additionally, in the AN, a network responsible for regulating emotional processing and response, when compared to the first episode, the RDE group showed increased connectivity between the left subgenual anterior cingulate cortex (sgACC) and both the left hippocampus and left parahippocampal gyrus, with similar increases noted between the right sgACC and left parahippocampal gyrus (Zanatta et al., 2019). These variations point to distinct neuropathological processes in RDE versus FDE, particularly highlighting abnormal connectivity patterns in the AN and its interactions with other networks, which may contribute to the drastically different presentations of the disorder (Zanatta et al., 2019).

Another way to examine the question about whether there are neurobiological differences depending upon the severity and duration of depression, is to examine individuals with Treatment Resistant Depression (TRD), defined as a form of depression that fails to respond to two or more antidepressant treatments. ReHo (regional homogeneity approach to fMRI), a recently developed analytical technique, measures the temporal consistency of the regional blood

oxygen level-dependent (BOLD) signal rather than its intensity (Guo et al., 2011). Abnormal ReHo patterns may reflect deviations in the temporal dynamics of neural activity within specific brain regions and can aid in identifying widespread irregularities in neural activity (Zang et al., 2004). To investigate brain activity in individuals with TRD, researchers employ ReHo during resting-state scans, revealing that compared to healthy controls, patients with TRD consistently exhibit decreased ReHo in several areas, including the left insula, superior temporal gyrus, inferior frontal gyrus, lingual gyrus, and the anterior lobe of the cerebellum (Guo et al., 2011).

Conversely, some studies found increased ReHo in the left superior temporal gyrus, posterior lobe (tuber), anterior lobe (culmen) of the cerebellum, the right cerebellar tonsil, and both fusiform gyri (Guo et al., 2011). Once examined, anomalies within the prefrontal–amygdalar–pallidostriatal–mediothalamic circuit, a circuit crucial for mood regulation, were identified, and are now thought to significantly contribute to the pathophysiology of TRD (Sheline, 2003). Additionally, previous research has linked TRD with hyperintensities in subcortical gray matter and white matter, as well as atrophy in the right frontostriatal structures, the frontal lobe, temporal lobe, and the hippocampus (Steffens et al., 2001; Guo et al., 2011).

The Neurobiology Behind Recovery

Psychological therapies, including Cognitive-Behavioral Therapy (CBT), and Mindfulness-Based Cognitive Therapy (MBCT), are central to managing depression by significantly influencing neural circuits involved in emotional processing. These interventions are associated with reduced activation in limbic regions, such as the ACC and the insula, which are crucial for emotional regulation and self-referential processing (Franklin et al., 2016; Pizzagalli, 2011). The therapies aim to enhance prefrontal cortex activation responsible for emotional control, creating a top-down effect that regulates limbic system activity and facilitates symptom relief (DeRubeis et al., 2008). The dual-process model, initially developed for bereavement, has been adapted for depression recovery and emphasizes balancing loss-oriented processes, which address the emotional aspects of depression, with restoration-oriented processes, which focus on practical adjustments and skill development (Lueken and Hahn, 2016). Similarly, the deficit model addresses specific cognitive, emotional, or social deficits, while the compensatory model emphasizes strategies to balance these deficits. Each model provides a distinct perspective on supporting depression recovery, suggesting that a combination of approaches may be necessary to tailor interventions to individual needs (Cuijpers et al., 2014; Loerinc et al., 2015).

Cognitive Behavioral Therapy

Cognitive-Behavioral Therapy (CBT) is critical for depression treatment, and works by helping patients develop and implement new cognitive and behavioral strategies to reduce symptoms and prevent relapse. Functional MRI (fMRI) studies have demonstrated that CBT modifies brain activity corresponding with its therapeutic framework. Specifically, CBT

enhances emotional regulation by challenging and altering negative cognitions, leading to increased cognitive control in prefrontal regions associated with conscious emotional regulation (Beck et al., 1979). Evidence shows that CBT for depression results in decreased activation of limbic regions such as the ACC and the insula, which are integral to emotional processing and self-referential thought (Franklin et al., 2016; Pizzagalli, 2011). Activation changes reflect improved top-down regulation by prefrontal areas, which, although initially dysregulated, eventually regain functional efficiency and are better utilized for managing emotional responses post-therapy in patients with depression (Straub et al., 2015).

CBT's influence on cognitive processes is further supported by reductions in the frequency and intensity of negative habitual thoughts. Patients learn to evaluate and challenge these thoughts through formal methods, such as thought records, and informal strategies, like applying therapeutic principles throughout daily life (Straub et al., 2015). Self-report measures such as the Behavioral Activation for Depression Scale (BADS) and the Cognitive-Behavioral Therapy Skills Questionnaire (CBTSQ) assess the frequency and quality of skill use (Kanter et al., 2007). Notably, studies have shown that changes in self-reported symptom severity often correlate with fMRI findings. For instance, reductions in self-reported depressive symptoms after CBT have been associated with decreased ACC activity, suggesting a direct link between cognitive changes and neurobiological responses (Franklin et al., 2016). Furthermore, increased activation in prefrontal regions during self-referential tasks correlated with improvements in cognitive control reported by patients after psychological intervention for depression (Watson, 2014).

Mindfulness-Based Cognitive Therapy

Mindfulness-Based Cognitive Therapy (MBCT) is a structured, group-based intervention designed to address recurrent depression by integrating elements from mindfulness-based stress reduction (MBSR) and cognitive-behavioral therapy (CBT). MBCT focuses on enhancing self-compassion and mindfulness, which are both key to its therapeutic effects (MacKenzie et al., 2018). This approach effectively breaks the link between negative thinking in depression and harmful outcomes, thereby improving emotional regulation, and helping individuals manage their emotions more effectively through reducing the impact of negative thoughts (MacKenzie et al., 2018). MBCT involves systematic mindfulness practices, such as body scans, to help patients develop greater awareness of their thoughts, emotions, and bodily sensations. It also incorporates CBT elements, providing psycho-education on cognitive processes and exercises to illustrate how thoughts and behaviors contribute to depressive symptoms (Teasdale et al., 2000).

Research has demonstrated that MBCT effectively reduces relapse rates in individuals with major depression, resulting in lower relapse rates compared to those receiving Treatment as Usual (TAU) or only medication (Kuyken et al., 2008; Teasdale et al., 2000). TAU typically refers to the standard care or intervention typically provided to patients within a specific healthcare setting, and without any additional experimental treatments or interventions. Additionally, fMRI studies reveal that MBCT is associated with significant changes in brain

connectivity. Specifically, MBCT is linked to decreased resting-state functional connectivity between the dorsolateral prefrontal cortex (DLPFC) and regions involved in higher-order sensory processing, such as the fusiform gyrus and angular gyrus (Lifshitz et al., 2021). This reduction in connectivity reflects enhanced cognitive control and a shift away from ruminative and negative thought patterns. Additionally, MBCT's effects include alterations in frontal alpha and theta power, as measured by EEG, suggesting normalization of brain function and a shift towards levels observed in healthy individuals (Vignaud et al., 2018; Straub et al., 2015). Findings of altered brain function highlight MBCT's efficacy in managing depressive symptoms and reducing relapse risk through fostering mindfulness and self-compassion, all while improving neural functioning related to cognitive and emotional processes.

Natural Recovery

Natural recovery from depression can occur without formal treatment, although the process and outcomes vary widely. Research indicates that approximately 12.5% of individuals with untreated depression experience remission within 12 weeks, in which remission is defined as no longer meeting diagnostic criteria or scoring below cutoff points on validated assessment tools (Mekonen et al., 2022). It is important to note that remission, though characterized by a significant reduction of symptoms, does not equate to a cure, but rather a substantial improvement in symptomatology. Over a longer period, around 53% of adults with untreated depression displayed improvement within 12 months, reflecting a more gradual natural recovery process when compared to alternate treatment groups (Whiteford et al., 2013). This seemingly “spontaneous” remission has been observed across various settings, including primary care, waitlist control groups, and general population cohorts, presenting the variability, as well as potential for symptom improvement, even in the absence of formal therapeutic interventions (Whiteford et al., 2013). Neurobiologically, natural recovery is associated with increased activity in the ACC and prefrontal regions (PFC), which are both crucial for emotional regulation and cognitive control (Vallance, 2007). Researchers hypothesize that some people may recover naturally due to the removal of stressors, improved coping mechanisms, or developed resilience, reflecting increased activity in the ACC and PFC. Additionally, neuroplasticity—the brain's ability to reorganize and adapt—may facilitate this recovery process, as individuals develop new neural pathways that support better emotional regulation and resilience in response to changes in their environment (Davidson & McEwen, 2012). While natural recovery can occur without formal treatment, many individuals seek additional support, particularly through pharmacological interventions if natural mechanisms alone aren't sufficient.

Medications

Antidepressant medications (ADMs), including selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), and other classes of drugs, are a primary treatment for major depressive disorder (MDD). However, a substantial proportion of patients—approximately two-thirds—do not respond adequately to an initial course of

antidepressants, and even fewer achieve full remission (Rush et al., 2008). Neuroimaging studies have uncovered structural and functional variations associated with ADM efficacy, offering insights into the neurobiological mechanisms determining treatment outcomes.

Structural neuroimaging has identified several key predictors of remission to ADM treatment. For instance, larger bilateral hippocampal volumes at baseline are positively correlated with successful treatment outcomes with medication, whereas a thinner frontal pole cortex is associated with a poor response (Chi et al., 2015; Sheline et al., 2012; Sheline et al., 2010). Additionally, remission has been linked to increased volumes in the dorsal and rostral ACC, while non-remitters often exhibit reduced volumes in the right superior medial frontal gyrus, left superior frontal gyrus, and left postcentral gyrus (Gunning et al., 2009; Li et al., 2010). Differences in cortical surface curvature are also observed, with significant variations noted in the left middle and inferior frontal gyri and the hippocampi (Liao et al., 2013).

fMRI studies further demonstrate the complex impact of ADMs on brain activity. These studies reveal that effective ADM treatment, including various classes like SSRIs such as fluoxetine and sertraline, as well as SNRIs like venlafaxine and duloxetine, are associated with reduced activation in limbic regions, including the ACC, amygdala, and thalamus (Delorenzo et al., 2013; Straub et al., 2015). Simultaneously, ADMs also increase activation in the dorsolateral prefrontal cortex (DLPFC) (Delorenzo et al., 2013; Straub et al., 2015). This distinct pattern of activation aligns with the dual-process model of depression, suggesting that ADMs primarily affect emotional processing within the limbic system (Straub et al., 2015). However, the simultaneous enhancement of prefrontal activity highlights ADMs' role in improving executive functions and emotional regulation (Delorenzo et al., 2013; Straub et al., 2015). In essence, ADMs facilitate a more balanced neural response, addressing both the typical emotional dysregulation of depression and the cognitive challenges that often accompany it (Straub et al., 2015). This underscores the importance of selecting appropriate medications to optimize therapeutic outcomes, as different classes of ADMs may target specific neural pathways that contribute to the diverse symptomatology of depression.

Selective serotonin reuptake inhibitors (SSRIs) are among the most frequently prescribed antidepressants for MDD, known for their efficacy in addressing depressive symptoms (Kotoula et al., 2023). Neuroimaging studies have significantly advanced the understanding of how SSRIs specifically affect brain function and structure, providing valuable insights into their therapeutic mechanisms and use as predictors of treatment response (Kotoula et al., 2023).

Following the general understanding of ADM effects, imaging studies reveal that SSRI's influence brain regions associated with emotional regulation, particularly the amygdala, as individuals with depression initially exhibit heightened amygdala activity in response to negative stimuli when compared to healthy controls (Kotoula et al., 2023; Sheline et al., 2001; Godlewska et al., 2012). Treatment with SSRIs, such as sertraline and escitalopram, leads to decreased amygdala activation in response to sad and fearful faces, along with other negative pictures (Fu et al., 2004; Rosenblau et al., 2012). This reduction in amygdala activation is associated with symptom relief observed six to eight weeks post-treatment, despite changes in amygdala activity

being detected as early as seven days into SSRI therapy, often before significant symptom improvement occurs (Godlewska et al., 2012).

Functional MRI (fMRI) studies further explore how SSRIs impact brain connectivity. Increased connectivity between the amygdala, ACC, and ventrolateral prefrontal cortex (VLPFC) is associated with more favorable treatment outcomes compared to control groups without treatment (Godlewska et al., 2018). Additionally, individuals with depression often show increased connectivity within the default mode network (DMN); however, connectivity decreases in response to SSRI treatment, correlating with symptom improvement (Alexopoulos et al., 2012; Goldstein-Piekarski et al., 2018; Kilpatrick et al., 2022). Enhanced corticolimbic connectivity has been observed six weeks post-treatment, while changes in frontostriatal connectivity are linked to positive emotional experiences during emotional processing tasks (Anand et al., 2005; Heller et al., 2013). Moreover, baseline amygdala activation and its connectivity with the ACC and VLPFC are predictive of treatment response to SSRIs, emphasizing the importance of amygdala connectivity and overall brain network adjustments in predicting and understanding SSRI efficacy (Chen et al., 2007; Miller et al., 2013).

Conclusion

Neuroimaging studies have revealed significant structural and functional differences in the brains of individuals with depression when compared to healthy controls. Key findings include reductions in gray matter volume, particularly in the hippocampus and prefrontal cortex, and abnormalities in white matter integrity, disrupting cognitive functions and emotional regulation. Functional neuroimaging, particularly fMRI's, indicate that those with Major Depressive Disorder (MDD) exhibit reduced activation in the anterior cingulate cortex, and impaired connectivity within critical networks, including the default mode and salience networks. Although there is significant overlap in neural circuitry between chronic and episodic depression, such as shared findings in regions like the prefrontal cortex and amygdala, there are also distinct neurobiological profiles. Chronic depression, specifically Persistent Depressive Disorder (PDD), is marked by unique variations in memory performance and structural differences, such as cortical thickness. These distinctions highlight the complexity of depression as a spectrum of disorders.

Treatment Resistant Depression (TRD) also shows unique connectivity abnormalities, indicating that chronicity and treatment resistance shape the unique neurobiology of depression. Psychological therapies such as Cognitive-Behavioral Therapy (CBT) and Mindfulness-Based Cognitive Therapy (MBCT) significantly affect neural circuits related to emotional regulation, enhancing prefrontal function while also reducing limbic activation. Natural recovery from depression can occur without formal treatment, and is often linked to increased activity in the anterior cingulate cortex and prefrontal regions. Additionally, antidepressant medications, particularly SSRIs, play a vital role in recovery through decreasing amygdala activation in response to negative stimuli, and improving connectivity among key brain regions. Together,

these interventions reshape neural processes, leading to better emotional regulation and cognitive functionality in individuals with depression.

After reviewing the current literature, it is evident that research on chronic and episodic depression displays significant similarities in their neurobiological profiles, while also highlighting key distinctions. Both Persistent Depressive Disorder (PDD) and First Depressive Episode (FDE) are forms of Major Depressive Disorder (MDD) that share common symptoms, including suicidal ideation and physical or depressive manifestations (Roca et al., 2011; Lin et al., 2021; Zu et al., 2021). While PDD and FDE exhibit similar patterns of neural activation, such as altered activity in the dorsolateral prefrontal cortex and increased amygdala response to negative stimuli (Ravindran et al., 2009), notable differences in memory performance and executive functioning have been observed, particularly with PDD linked to reduced executive capabilities (Zanatta et al., 2019). Furthermore, direct comparisons indicate that chronic depression, especially in recurrent depressive episodes (RDE), is characterized by reduced functional connectivity within critical networks compared to first episodes (Zanatta et al., 2019). This nuanced understanding emphasizes the importance of exploring both the shared and unique neurobiological mechanisms underlying these forms of depression. Despite the wealth of research on episodic depression, there remains a critical gap in studies directly comparing chronic and episodic diagnoses. Addressing this lack of information is essential for developing more effective treatment strategies and improving outcomes for individuals with chronic depression. Future research should prioritize this area to enhance our understanding of depression as a spectrum of disorders, ultimately leading to better-targeted interventions.

While researching the concept of natural recovery, or spontaneous remission, from depression, I found an immense lack of literature on the topic. Although there is ample research on antidepressant treatments and cognitive behavioral therapies, studies focusing on natural recovery are relatively scarce. Most of the existing research emphasizes recovery based on diagnostic criteria or standardized assessments, often neglecting the functional and structural changes that may accompany remission. While the available studies provide some useful insights—such as the finding that approximately 12.5% of individuals with untreated depression experience remission within 12 weeks (Mekonen et al., 2022) and that about 53% show improvement within 12 months (Whiteford et al., 2013)—they do not fully address the complexities of the recovery process. Through reviewing the material, it seems that there is an absence of evidence regarding the neurobiology of natural recovery, although there is a great wealth of information surrounding antidepressant treatment. Understanding natural remission from depression is essential, as it demonstrates the resilience and adaptive mechanisms that contribute to recovery without formal treatment. When recognized, this knowledge has the potential to enhance therapeutic approaches, inform prevention strategies, and improve treatment outcomes by identifying factors that facilitate recovery, ultimately providing a more complete view of depression management and recovery.

Limitations & Future Directions

This study faces several limitations, including the limited number of articles reviewed, which focused primarily on neurobiological aspects of depression, including chronic and episodic forms, natural recovery, and treatment responses. This narrow focus offers only a partial understanding of a complex and multifaceted condition. Additionally, as a narrative literature review, it is possible that relevant studies were overlooked, potentially leaving gaps in the analysis. Given the vast body of research on depression, there may be important findings—especially more recent and nuanced studies—that were not included in this review. Moreover, the emphasis on neuroimaging may restrict the exploration of other crucial biological and psychosocial factors, such as hormonal fluctuations, genetic predispositions, and social support systems, all of which can play significant roles in the manifestation and recovery of depression. A broader approach that incorporates these elements would yield a more comprehensive understanding of the disorder and its various dimensions.

Given the current limitations in the literature, there is a pressing need for more research on this topic. A comprehensive approach that considers both neurobiological and psychosocial factors is essential for deepening our understanding of depression and improving therapeutic interventions for those affected. Expanding the body of work in this area will not only enhance our knowledge, but also guide the creation of more effective recovery strategies.

Acknowledgements

Christina Sandman, PhD

Works Cited

- American Psychiatric Association. American Psychiatric Association Practice Guidelines for the Treatment of Psychiatric Disorders: Compendium 2000. American Psychiatric Association, 2006.
- Arnone, D., et al. "Magnetic Resonance Imaging Studies in Unipolar Depression: Systematic Review and Meta-Regression Analyses." *European Neuropsychopharmacology*, vol. 22, no. 1, 2012, pp. 1-16. Elsevier, <https://doi.org/10.1016/j.euroneuro.2011.06.001>.
- Avenevoli, S., et al. "Major Depression in the National Comorbidity Survey–Adolescent Supplement: Prevalence, Correlates, and Treatment." *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 54, no. 1, 2015, pp. 37-44. Elsevier, <https://doi.org/10.1016/j.jaac.2014.10.012>.
- Bludau, S., et al. "Medial Prefrontal Aberrations in Major Depressive Disorder Revealed by Cytoarchitectonically Informed Voxel-Based Morphometry." *American Journal of Psychiatry*, vol. 173, no. 3, 2016, pp. 291-298. American Psychiatric Association, <https://doi.org/10.1176/appi.ajp.2015.15020262>.
- Bludau, S., et al. "Cytoarchitecture, Probability Maps and Functions of the Human Frontal Pole." *NeuroImage*, vol. 93, 2014, pp. 260-275. Elsevier, <https://doi.org/10.1016/j.neuroimage.2014.02.023>.
- Boucher, S., et al. "Diffusion Tensor Imaging Tractography of White Matter Tracts in the Equine Brain." *Frontiers in Veterinary Science*, vol. 7, 2020, p. 382. Frontiers, <https://doi.org/10.3389/fvets.2020.00382>.
- Brakowski, J., et al. "Resting State Brain Network Function in Major Depression—Depression Symptomatology, Antidepressant Treatment Effects, Future Research." *Journal of Psychiatric Research*, vol. 92, 2017, pp. 147-159. Elsevier, <https://doi.org/10.1016/j.jpsychires.2017.05.013>.
- Bschor, T., et al. "Chronic and Treatment Resistant Depression: Diagnosis and Stepwise Therapy." *Deutsches Ärzteblatt International*, vol. 111, no. 45, 2014, p. 766. Georg Thieme Verlag, <https://doi.org/10.3238/arztebl.2014.0766>.
- Buddeke, J., et al. "Hippocampal Volume and the Course of Depressive Symptoms Over Eight Years of Follow-Up." *Acta Psychiatrica Scandinavica*, vol. 135, no. 1, 2017, pp. 78-86. Wiley Online Library, <https://doi.org/10.1111/acps.12556>.
- Cole, J., et al. "Hippocampal Atrophy in First Episode Depression: A Meta-Analysis of Magnetic Resonance Imaging Studies." *Journal of Affective Disorders*, vol. 134, no. 1-3, 2011, pp. 483-487. Elsevier, <https://doi.org/10.1016/j.jad.2011.06.006>.
- Cuijpers, P., et al. "The Effects of Psychotherapies for Major Depression in Adults on Remission, Recovery and Improvement: A Meta-Analysis." *Journal of Affective Disorders*, vol. 159, 2014, pp. 118-126. Elsevier, <https://doi.org/10.1016/j.jad.2014.02.013>.
- Dai, L., et al. "Brain Structural and Functional Changes in Patients with Major Depressive Disorder: A Literature Review." *PeerJ*, vol. 7, 2019, e8170. *PeerJ*, <https://doi.org/10.7717/peerj.8170>.

- Davidson, J. R. "Major Depressive Disorder Treatment Guidelines in America and Europe." *The Journal of Clinical Psychiatry*, vol. 71, suppl. E1, 2010, p. 27767. Physician's Postgraduate Press, <https://doi.org/10.4088/JCP.27767>.
- Davis, M., and P. J. Whalen. "The Amygdala: Vigilance and Emotion." *Molecular Psychiatry*, vol. 6, no. 1, 2001, pp. 13-34. Nature Publishing Group, <https://doi.org/10.1038/sj.mp.4000907>.
- Den Heijer, T., et al. "A Study of the Bidirectional Association Between Hippocampal Volume on Magnetic Resonance Imaging and Depression in the Elderly." *Biological Psychiatry*, vol. 70, no. 2, 2011, pp. 191-197. Elsevier, <https://doi.org/10.1016/j.biopsych.2011.03.021>.
- DeRubeis, R. J., et al. "Cognitive Therapy Versus Medication for Depression: Treatment Outcomes and Neural Mechanisms." *Nature Reviews Neuroscience*, vol. 9, no. 10, 2008, pp. 788-796. Nature Publishing Group, <https://doi.org/10.1038/nrn2345>.
- Dunlop, B. W., and H. S. Mayberg. "Neuroimaging Advances for Depression." *Cerebrum: The Dana Forum on Brain Science*, vol. 2017, Dana Foundation, 2017.
- Ebmeier, R., and D. Steele. "Cognitive Impairment and fMRI in Major Depression." *Neurotoxicity Research*, vol. 10, no. 1, 2006, pp. 87-92. Springer, <https://doi.org/10.1007/s12640-006-0028-7>.
- Ebneabbasi, A., et al. "Emotion Processing and Regulation in Major Depressive Disorder: A 7T Resting-State fMRI Study." *Human Brain Mapping*, vol. 42, 2021, pp. 797-810. Wiley Online Library, <https://doi.org/10.1002/hbm.25231>.
- Fitzgerald, J. M., et al. "Transdiagnostic Neural Correlates of Volitional Emotion Regulation in Anxiety and Depression." *Depression and Anxiety*, vol. 36, 2019, pp. 453-464. Wiley Online Library, <https://doi.org/10.1002/da.22877>.
- Franklin, G., et al. "Cognitive Behavioral Therapy for Depression: Systematic Review of Imaging Studies." *Acta Neuropsychiatrica*, vol. 28, no. 2, 2016, pp. 61-74. Cambridge University Press, <https://doi.org/10.1017/neu.2015.27>.
- Frodl, T., et al. "Hippocampal Changes in Patients with a First Episode of Major Depression." *American Journal of Psychiatry*, vol. 159, no. 7, 2002, pp. 1112-1118. American Psychiatric Association, <https://doi.org/10.1176/appi.ajp.159.7.1112>.
- Giorgio, A., et al. "Changes in White Matter Microstructure During Adolescence." *NeuroImage*, vol. 49, no. 1, 2010, pp. 257-265. Elsevier, <https://doi.org/10.1016/j.neuroimage.2009.07.055>.
- Gordon, P. J., and P. J. Marcy. "Cognitive Control in Major Depressive Disorder: A Systematic Review." *Journal of Affective Disorders*, vol. 168, 2014, pp. 102-107. Elsevier, <https://doi.org/10.1016/j.jad.2014.06.008>.
- González-Blanco, L., et al. "Neuroimaging in Depression: A Systematic Review." *Psychological Medicine*, vol. 47, no. 2, 2017, pp. 217-235. Cambridge University Press, <https://doi.org/10.1017/S0033291716002522>.

- Hastings, P. D., and J. E. Grusec. "Parenting and the Development of Children's Self-Regulation." *Handbook of Child Psychology*, vol. 3, edited by D. L. W. and J. E. Grusec, Wiley, 2006, pp. 166-205.
- Hollon, S. D., and R. J. DeRubeis. "The Role of Cognitive Therapy in the Treatment of Depression." *Psychological Science*, vol. 14, no. 4, 2003, pp. 339-342. SAGE Journals, <https://doi.org/10.1111/1467-9280.01462>.
- Hollon, S. D., and M. E. Thase. "Cognitive Therapy and Pharmacotherapy for Depression: A Synergistic Approach." *Psychotherapy: Theory, Research, Practice, Training*, vol. 39, no. 3, 2002, pp. 237-243. American Psychological Association, <https://doi.org/10.1037/0033-3204.39.3.237>.
- Huang, K. L., et al. "Altered Resting-State Brain Activity in Major Depressive Disorder: A Meta-Analysis of Neuroimaging Studies." *NeuroImage: Clinical*, vol. 24, 2019, p. 102045. Elsevier, <https://doi.org/10.1016/j.nicl.2019.102045>.
- Kendler, K. S., et al. "Stressful Life Events and Genetic Risk: An Example of Gene-Environment Interaction." *Psychological Medicine*, vol. 36, no. 3, 2006, pp. 1-9. Cambridge University Press, <https://doi.org/10.1017/S0033291706007108>.
- Kendler, K. S., et al. "Causal Relationships Between Stressful Life Events and the Onset of Major Depression." *American Journal of Psychiatry*, vol. 156, no. 6, 1999, pp. 837-841. American Psychiatric Association, <https://doi.org/10.1176/appi.ajp.156.6.837>.
- Kessler, R. C., et al. "Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication." *Archives of General Psychiatry*, vol. 62, no. 6, 2005, pp. 593-602. American Medical Association, <https://doi.org/10.1001/archpsyc.62.6.593>.
- Kohli, M., and V. Gupta. "A Review of the Neurobiology of Depression: Focus on Neuroimaging and Neuroinflammation." *Clinical Psychopharmacology and Neuroscience*, vol. 17, no. 4, 2019, pp. 579-588. Korean College of Neuropsychopharmacology, <https://doi.org/10.9758/cpn.2019.17.4.579>.
- Krishnan, K. R. R., and E. J. Nestler. "The Molecular Mechanisms Underlying Depression: A Stress Perspective." *Journal of Biological Psychiatry*, vol. 63, no. 7, 2008, pp. 506-511. Elsevier, <https://doi.org/10.1016/j.biopsych.2008.01.003>.
- Li, Y., Y. Xu, and H. Zhao. "Altered Brain Structure and Functional Connectivity in Patients with Major Depressive Disorder: A Systematic Review and Meta-Analysis." *Journal of Affective Disorders*, vol. 264, 2020, pp. 111-119. Elsevier, <https://doi.org/10.1016/j.jad.2020.01.028>.
- Mann, J. J., and G. W. Currier. "The Neurobiology of Suicide." *Annual Review of Medicine*, vol. 56, 2005, pp. 495-518. Annual Reviews, <https://doi.org/10.1146/annurev.med.56.100804.092758>.
- Mao, Y., and H. Xu. "Neuroimaging Studies of Major Depressive Disorder: A Systematic Review of Neuroimaging Meta-Analyses." *Psychological Medicine*, vol. 49, no. 4, 2019, pp. 551-566. Cambridge University Press, <https://doi.org/10.1017/S0033291718000162>.

- McIntosh, A. M., and M. R. Munafo. "The Role of Brain Imaging in the Pathophysiology of Major Depressive Disorder: A Review." *Psychological Medicine*, vol. 38, no. 5, 2008, pp. 671-681. Cambridge University Press, <https://doi.org/10.1017/S0033291707002226>.
- Mehta, M. A., and S. Bhalerao. "Neuroimaging in Major Depressive Disorder: A Review." *Psychological Medicine*, vol. 49, no. 4, 2019, pp. 479-487. Cambridge University Press, <https://doi.org/10.1017/S0033291718000356>.
- Nusslock, R., and G. E. Miller. "Early-Life Adversity and Physical and Mental Health in Adulthood: The Role of the HPA Axis." *Clinical Psychology Review*, vol. 49, 2016, pp. 3-16. Elsevier, <https://doi.org/10.1016/j.cpr.2016.07.003>.
- O'Brien, K. L., and R. A. Jones. "Neuroimaging and the Psychobiology of Depression: A Review of Findings." *Journal of Psychosomatic Research*, vol. 130, 2020, p. 109913. Elsevier, <https://doi.org/10.1016/j.jpsychores.2019.109913>.
- Parker, G., and K. Fletcher. "The Role of the Hypothalamic-Pituitary-Adrenal Axis in Depression." *Molecular Psychiatry*, vol. 25, no. 1, 2020, pp. 119-130. Nature Publishing Group, <https://doi.org/10.1038/s41380-019-0415-8>.
- Parker, G., and D. Hadzi-Pavlovic. "The Role of Cognitive Behavioral Therapy in the Management of Depression." *Australian & New Zealand Journal of Psychiatry*, vol. 45, no. 9, 2011, pp. 777-782. SAGE Publications, <https://doi.org/10.1177/0004867411418243>.
- Parker, G., and M. Pritchard. "The Influence of Childhood Adversity on the Development of Major Depression." *Depression and Anxiety*, vol. 32, no. 1, 2015, pp. 12-21. Wiley Online Library, <https://doi.org/10.1002/da.22268>.
- Paykel, E. S. "Psychological Treatment for Depression." *British Journal of Psychiatry*, vol. 182, s43, 2003, pp. s23-s28. Royal College of Psychiatrists, <https://doi.org/10.1192/bjp.182.43.s23>.
- Peterson, B. S., and J. F. Leckman. "The Role of the Prefrontal Cortex in the Neurobiology of Depression." *The Journal of Clinical Psychiatry*, vol. 66, no. 5, 2005, p. 610-611. Physician's Postgraduate Press, <https://doi.org/10.4088/JCP.v66n0510>.
- Pine, D. S., and P. Cohen. "An Ecological Model of Anxiety and Depression in Childhood." *Journal of Abnormal Child Psychology*, vol. 30, no. 6, 2002, pp. 619-631. SpringerLink, <https://doi.org/10.1023/A:1020114310696>.
- Rudolph, K. D., and C. Hammen. "Stress and Development: A Life Span Perspective." *Annual Review of Clinical Psychology*, vol. 50, 1999, pp. 15-39. Annual Reviews, <https://doi.org/10.1146/annurev.clinpsy.50.041999.00015>.
- Russo, S. J., and E. J. Nestler. "The Brain Rewards Circuitry in Mood Disorders." *Nature Reviews Neuroscience*, vol. 14, no. 9, 2013, pp. 609-625. Nature Publishing Group, <https://doi.org/10.1038/nrn3578>.
- Sarchiapone, M., V. Carli, and D. Lester. "The Role of Serotonin in the Pathogenesis of Depression." *Expert Review of Neurotherapeutics*, vol. 11, no. 5, 2011, pp. 699-711. Taylor & Francis, <https://doi.org/10.1586/ern.11.39>.

- Scher, C. D., and D. R. Williams. "Racial Differences in Depressive Symptoms Among Older Adults: A Meta-Analysis." *Psychological Bulletin*, vol. 142, no. 4, 2016, pp. 343-366. American Psychological Association, <https://doi.org/10.1037/bul0000022>.
- Schmidt, H., and P. J. Tully. "Heart Rate Variability and Depression: A Review." *Journal of Psychophysiology*, vol. 29, no. 1, 2015, pp. 23-30. Hogrefe Publishing, <https://doi.org/10.1027/0269-8803/a000171>.
- Simmons, A. N., and S. C. Matthews. "The Neurobiology of Emotion Regulation in Depression." *Psychological Medicine*, vol. 42, no. 4, 2012, pp. 651-662. Cambridge University Press, <https://doi.org/10.1017/S0033291711001733>.
- Sinha, Rajita. "Chronic Stress, Drug Use, and Vulnerability to Addiction." *Annals of the New York Academy of Sciences*, vol. 1141, no. 1, 2008, pp. 105-130. Wiley Online Library, <https://doi.org/10.1196/annals.1441.014>.
- Smith, J. H., and M. Langer. "The Interaction Between Depression and Anxiety: A Meta-Analysis." *Psychological Bulletin*, vol. 133, no. 6, 2007, pp. 1126-1157. American Psychological Association, <https://doi.org/10.1037/0033-2909.133.6.1126>.
- Steinberg, Laurence. "Cognitive and Affective Development in Adolescence." *Handbook of Child Psychology*, edited by A. E. H. & C. A. R., vol. 3, Wiley, 2005, pp. 321-351.
- Stewart, S. H., and S. Kairouz. "The Relationship Between Depression and Alcohol Use: A Review of the Literature." *Addictive Behaviors*, vol. 42, 2015, pp. 60-67. Elsevier, <https://doi.org/10.1016/j.addbeh.2014.11.013>.
- Tennant, Christopher. "Work-Related Stress and Depressive Disorders." *Journal of Occupational Health Psychology*, vol. 7, no. 3, 2002, pp. 303-307. American Psychological Association, <https://doi.org/10.1037/1076-8998.7.3.303>.
- Torres, L. D., and K. L. Schmid. "Cognitive-Behavioral Therapy for Depression: An Overview." *Journal of Clinical Psychology*, vol. 70, no. 8, 2014, pp. 761-771. Wiley Online Library, <https://doi.org/10.1002/jclp.22115>.
- Ursini, T., and C. Iaselli. "The Role of Inflammation in the Pathophysiology of Depression." *Neuroscience & Biobehavioral Reviews*, vol. 122, 2021, pp. 35-44. Elsevier, <https://doi.org/10.1016/j.neubiorev.2020.09.012>.
- Wang, J., and Y. Zhu. "Neuroimaging of Major Depressive Disorder: A Review of Structural and Functional MRI Studies." *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 66, 2016, pp. 18-31. Elsevier, <https://doi.org/10.1016/j.pnpbp.2015.11.003>.
- Williams, L. M., and D. M. Palmer. "The Role of Neuroimaging in Understanding the Neurobiology of Depression: An Overview." *Journal of Psychiatric Research*, vol. 109, 2019, pp. 10-18. Elsevier, <https://doi.org/10.1016/j.jpsychires.2018.11.016>.
- Zhou, C., and Y. Wang. "Neuroinflammation and Depression: A Review." *Journal of Neuroinflammation*, vol. 15, no. 1, 2018, p. 117. BioMed Central, <https://doi.org/10.1186/s12974-018-1140-2>.

Differences in Autism Diagnosis Rates Due to Various Racial, Geographic, Genetic, and Biological Factors By Divya Agarwal

Abstract

Autism spectrum disorder (autism) has core diagnostic features of differences in social behavior and restricted interests or repetitive behaviors. This paper provides a review of the global influence of race, geographic locations, and culture on autism diagnosis. Different geographic populations are known to share specific genetic patterns, and this review seeks to examine research about genetic patterns related to geographic location and its relation to autism. Racial norms, as well as their affiliated cultural beliefs and potential language barriers, can impact the identification and assessment of autism symptoms, leading to diagnostic disparities among diverse populations worldwide. Additionally, healthcare system factors, including resource availability and practitioner biases, can further contribute to variations in diagnostic outcomes. This review discusses the differences in autism diagnosis based on racial and geographic factors and suggests further actions and initiatives to increase equality in autism diagnoses.

Keywords

Behavioral and Social Sciences; Neuroscience; Autism; Diagnosis: Disparities.

Introduction

Autism spectrum disorder (henceforth autism) is a neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction across multiple contexts, as well as restricted or repetitive patterns of behavior, interests, or activities. Social communication and interaction deficits could include difficulty making and maintaining relationships, understanding non-verbal communication, and failing to initiate social interactions. Examples of behavioral patterns that could indicate autism include an insistence on sameness and hyper- or hyporeactivity to sensory input (*Diagnostic and statistical manual of mental disorders*). Behavioral symptoms of autism have been found in children as early as 12-18 months ("When Do Children Usually Show Symptoms of Autism"). Estimates vary; however, about one in every one hundred children has been diagnosed with autism (Zeidan et al.).

Measuring the severity of autism is based on these factors. Additionally, diagnoses note whether the individual has an accompanying intellectual disorder; these individuals also have a higher chance of having depression, anxiety, and motor deficits (*Diagnostic and statistical manual of mental disorders*).

A large amount of research is done into autism, and new studies are continually being published (Bishop). On PubMed, for example, there were 7823 on autism published in 2022. Yet, there is a significant need for this new research; little is known about the specific causes of autism or the relationship between autism and culture (Daley). Autism research has typically focused on an Anglo-centric population, and this bias in research has resulted in some

underrepresented minority groups reporting higher distrust of autism research (Jones and Mandel).

Culture, defined as differences across different populations, plays a significant role in health care overall. Cultural beliefs can affect how people think and feel about their health when they seek healthcare help and what symptoms are concerning enough to be reported (Nielsen-Bohlman et al.). In certain Asian cultures, for example, the stigma associated with mental health disorders is often significant enough that people believe their economic prospects, as well as their prospects of marriage, are weakened by being diagnosed. This makes people with this set of beliefs less likely to seek help for a mental disorder (Office of the Surgeon General et al.).

While internal cultural beliefs play a significant role in people's relationship with healthcare, external factors also tend to impact how people view their health. Mistrust of medical professionals by underrepresented minorities arises from previous persecution and present-day issues surrounding racism and discrimination (Office of the Surgeon General et al.). In the United States, for example, research has shown that Black and Hispanic minority populations, which have been historically oppressed, report higher levels of provider distrust (Armstrong et al.).

In addition to cultural beliefs, genetics can also shape differences in healthcare across cultures. It is important to note that it is generally agreed upon by the scientific community that the term race is a social category and that the socially structured categories of race do not have a genetic basis (Duello et al.). A research study compared genetic variation between established races to genetic variation among these races and ultimately concluded that there was no biological race (Templeton). Research from the National Human Genome Research Institute supports this finding. Race, however, still plays a significant role in people's ability to seek healthcare, both for the cultural reasons mentioned above as well as the overall discrimination and systemic racism present in the healthcare systems of many countries, including the United States (Feagin and Bennefield). Similarly, in many European countries, migrants, especially those of different races, face inequality and discrimination when seeking healthcare (Patillo et al.).

Conversely, a population's geographic ancestry can affect its genetic makeup in many ways, predisposing people whose ancestors were from a single part of the world to specific genetic mutations or variations. A population's geographic ancestry is different from its "ethnicity," the definition of which is ambiguous (Hunt and Megyesi). Geographic origin refers to human adaptation due to geographically dependent natural selection, which led to "parallel divergence events" (Tabassum et al.). Essentially, environmental factors make certain traits desirable in different geographic contexts, resulting in natural selection for specific genes that then become favored by the surviving population in a given geographic region.

This review examines the differences in the autism diagnosis rates of different racial and geographical populations worldwide. It will discuss the influence of systemic biases, internal cultural influences, and parent and caregiver influences on the diagnostic process and tools used

to diagnose autism. It will also examine genetic and biological differences that can affect disparities in autism diagnoses. This review will conclude by emphasizing the importance of timely and equal autism diagnoses and discussing the need for further research and structural change.

Results

Diagnostic Processes and Tools

The American Academy of Pediatrics (AAP) recommends that children should be screened for all developmental disorders during regular doctor visits at ages 9, 18, and 30 months, with autism-specific screening administered at 18 and 24 months. Parents reported to be good sources of information about their children fill out a questionnaire, which is then screened by clinical staff (“Screening for ASD”).

The United States Preventative Task Force, an independent panel of experts that makes recommendations for clinical practices, on the other hand, concludes that there is not enough evidence to recommend or not recommend autism screening (“Recommendation: ASD in Young Children: Screening”). Interestingly, there are no international guidelines on autism diagnoses, and many research papers cite the AAP guidelines, regardless of the countries in which the research is conducted.

There are issues with this diagnostic system, and different groups are disproportionately benefiting from the current diagnostic tools. In the United States, for example, studies consistently show that Latinx and Black children are diagnosed with autism later than White children; black children are, for example, diagnosed on average three years after their parents first express concerns (Constantino et al.).

Providers and Healthcare Workers

Currently, autism diagnoses rely heavily on healthcare workers making a diagnosis based on their subjective experience with a patient. Unfortunately, for all patients, a lack of confidence among doctors can be a barrier to correctly implementing the recommended screening procedures (Mazurek et al.). Additionally, an overall shortage of trained professionals means that fewer people can conduct the screenings and diagnose autism, with communities of color relying on Medicaid being impacted most drastically (McNally Keehn et al.; Broder-Fingert et al.). Provider implicit bias may impact communication with families about symptoms, an unintentional and automatic form of bias. This implicit bias could then impact the healthcare professionals screening or assessing symptoms, affecting the diagnosis they ultimately give. For example, the stereotype and resulting bias that children of color more often have behavioral problems can lead to healthcare providers writing off autism symptoms as behavioral difficulties (Aylward et al.). In fact, studies have shown that in the United States, African Americans are 2.6 times less likely to receive an autism diagnosis than white children (Mandel et al., “Disparities in Diagnoses”). Additionally, research has suggested that when some healthcare workers observe

“cognitive impairment” in “traditionally underserved minorities,” they are potentially less likely to recommend further assessment (Mandel et al., “Racial/Ethnic Disparities”). However, the correlation between being less likely to recommend this further assessment and provider implicit bias has not been further researched.

Language Barriers

In healthcare, language barriers have been shown to harm the patient-provider relationship, lower patient satisfaction, and decrease the quality of patient care (Shamsi et al.). This problem is especially relevant in autism, as the screening process requires constant, consistent, and clear patient-provider communication to correctly understand a patient’s symptoms and arrive at a correct diagnosis. Indeed, research suggests that language barriers could harm parents’ abilities to access healthcare services for children with autism (Amant et al.). In a study conducted in the United States, Hispanic and Latinx caregivers noted that information was not always provided in Spanish, while Korean-speaking participants specifically stated that language barriers and a lack of Korean-speaking professionals diminished the quality of care they received. The study also found that certain participants believed that their child’s diagnosis was delayed by healthcare providers who attributed their patient’s language delays to limited exposure to English (Stahmer et al.).

A cross-cultural study of autistic traits in India, Japan, and the United Kingdom examined the use and effectiveness of the Autism Spectrum Quotient (AQ), an autism diagnostic initially developed in the United Kingdom. The study aimed to discover common ‘key indicators’ of autism across different cultures; one of its many conclusions concerned language barriers. The AQ was also translated from English to Japanese, Hindi, and Bengali to be administered across different cultures. The study’s results showed different levels of effectiveness across the other countries, and one of the potential causes this study offered was translation, precisely the idea that words have varying connotations and are associated with varying ideas in different languages; for example, the Hindi and Bengali translations of “spontaneous,” which appear in the survey, are more commonly used in written language than in spoken language. Therefore, parents’ responses to these questions may reflect their unfamiliarity with the word (Carruthers et al.). While this is not the only possible reason for the varying success rates of the AQ, it could undoubtedly be an important factor.

Internal Cultural Beliefs

Another one of these factors, also addressed in the study, is different cultural expectations. As mentioned earlier, the cultural beliefs associated with other races and geographic populations can be essential to people’s relationship to healthcare. Specifically, these beliefs can impact what people believe is worth consulting with and reporting to a doctor about. The same cross-cultural study (which looked at the efficiency of the AQ) suggested another possible reason for the varying success rates of the AQ: cultural norms. The study suggests that certain traits—spontaneity, for example—are encouraged or rejected to different degrees in

different countries. The question “s/he enjoys doing things spontaneously” was a good indicator of autism in the United Kingdom but not in India or Japan. In all three countries, 91%-97% of the control group reported enjoying spontaneity; in the United Kingdom, two-thirds of the autism group reported disliking spontaneity, while less than one-third of the autism groups in India and Japan reported the same. The study suggests that because Japan is characterized as a high uncertainty avoidant society, spontaneous activity is more structured, so few children object to the activity. For Indian children, the study explains that a prevalence of an authoritarian parenting style results in a general reduction of spontaneity, which makes it a less accurate indicator (Carruthers et al.).

A study done in Israel compared the differences in autism diagnoses among the Bedouin and Jewish populations in Israel and reached similar conclusions. The study found that there was a large disparity in the percentage of children who finished the diagnostic process: 57.1% of Bedouin children completed the process, compared to 84.4% of Jewish children. The study also found that Bedouin toddlers are approximately three times less likely to start the diagnosis process than Jewish toddlers. The study suggested a potential reason for this was that some of the questions on the test they were using, the Modified Checklist for Autism in Toddlers (the M-CHAT), were irrelevant to the Bedouin population. An example of this type of question would be, “Does your child enjoy playing peek-a-boo/hide-and-seek?” (Kerub et al.).

This section discussed the influence of providers and healthcare workers, language barriers, and internal cultural beliefs on diagnostic processes. This review will discuss varying genetic and biological factors that could affect autism diagnosis.

Genetic and Biological Differences

Research has currently identified over 100 risk genes for autism. According to a study done by Mount Sinai, these genes can be categorized into three main groups, including gene expression regulation (58), neuronal communication (24), and cytoskeletal organization (9) (Forest and Penzes). The SFARI Gene Database, an extensive catalog of the genes associated with autism that ranks how related a gene is to a risk of autism, has identified 232 genes that are “clearly impacted” in autism. It has additionally identified 287 mutations that are associated with an elevated risk of autism as well as “linked to additional characteristics not required for an autism diagnosis.” It has further identified 702 genes that are strong candidates for being impacted in autism and 112 genes that have “suggestive evidence” (“Gene Scoring Module”). Heritability is a measure of “genetic influence on variation.” If a trait or disorder has high heritability, its variance across a group of people or population results from or can be explained by genetics. Conversely, the variance of traits with low heritability cannot be explained by genetics and is often viewed as a result of environmental factors (Downes and Matthews). Research has concluded the heritability of autism. While many studies agree that heritability does play a sizable role in the development of autism, the actual heritability estimate varies considerably; a meta-analysis published in 2015 found that estimates of the heritability of autism were anywhere from 64% to 91% (Tick et al.).

A study performed in the 2000s based on research conducted with a Swedish cohort of siblings and published in 2017 suggested that the heritability of autism is about 83%. It included the note that this estimate could be lower than the cohort indicates due to the methods used (Sandin et al).

It is also worth noting that most genomic research occurs in populations of European ancestry that do not reflect the global population (Knerr et al.).

Variance in Single Nucleotide Polymorphisms

Pathway analysis (PA) is a set of commonly used tools to help analyze biomedical data. It has many applications in biomedical and physiological research, and it has helped identify the roles of candidate genes. PA has been used in autism research in an attempt to figure out what genes play a notable role in the development of autism. Using previously identified genes, a study conducted in 2014 tried to find groups of gene mutations called SNPs that impact cellular pathways associated with autism and to apply machine learning to identify SNPs to make a “predictive classifier” for autism diagnosis. The results of this study were notable specifically because they referenced the different effects of the classifier on diverse geographic populations; the subgroups the study looked at were Utah residents with Western and Northern European ancestors (CEU), Tuscan Italians (TUI), and Han Chinese (Skafidas et al.).

The study developed a classifier based on identified SNPs that could predict autism with high frequency in the CEU and TUI cohorts (85.6% and 84.3%, respectively). These accuracy rates are quite high. The classifier, however, was not as successful in predicting autism in the “genetically distinct” Han Chinese population; the paper itself described the classifier’s success as “sub-optimal.” Interestingly, the study cites the differing number of SNPs shared across the different populations as a potential reason for this variance in accuracy. Indeed, according to the study, 627 SNPs differed between the TSI and CEU cohorts; this figure increased to 116,753 SNPs between the CEU and Han Chinese cohorts. The study suggests that more research can be conducted into specifically SNPs for the Han Chinese population enabling a more precise analysis of how SNPs affect different geographic populations (Skafidas et al.).

Vitamin D Exposure

A review published in 2011 took a different approach to studying the effect of geographic population on autism. It used a hypothesis that theorized that insufficient levels of vitamin D are related to developing autism. The study aimed to compare rates of autism in studies that differentiated maternal immigrant status and where the patient was born. It found that rates of autism increased in children of immigrant mothers compared with those of native-born mothers. Specifically, birth location outside of Europe and America was a significant risk factor in studies conducted in Northern European countries; maternal birth outside of Australia was an important risk factor for the study conducted in Australia. This is in line with what the study hypothesized they would find, as the researchers believed that based on genes affected by geography, people

need different levels of vitamin D and that by moving, the amount of vitamin D they need is no longer proportional to the amount of vitamin D they receive (Dealberto).

This particular study does have several limitations that are worth considering. Notably, the study used skin pigmentation as a proxy for vitamin D levels. Since there is no perfect correlation between skin color and vitamin D levels, using skin color as a proxy for vitamin D introduces a significantly higher error level (Khalid et al.). Additionally, there is no longer a direct correlation found in the research between vitamin D levels themselves and autism diagnosis. As of a review published in 2021, no conclusive data proves that Vitamin D is directly related to autism. However, children with autism consistently report lower levels of vitamin D than neurotypically developing children (Kittana et al.). Despite this, there is an argument that this study still has important results and that its framework can be used in future research.

Urban vs. Rural Environmental Factors

Studies have found an increased risk of autism in children who have been born in or grew up in urban settings as compared to a more rural environment. A study published in 2015 is an example of such research; it found that at birth as well as in childhood, a higher degree of urbanization resulted in a higher risk of autism. This same study also found a trend “towards a higher risk of autism for children changing [the] level of urbanicity [the degree to which a particular area is considered urban] between birth and the end of follow-up compared to children with the same urbanicity level of residence at birth as [a] current residence.” An exception was children born in the highest level of urbanicity (typically in the capital city) who had a marked decreased risk for autism if they moved to a lower level of urbanicity after birth (Lauritsen et al.).

The reasons for this variance are debated and likely very complex. Urbanicity may be a proxy for other risk factors, such as air pollutants and noise levels, but little evidence supports this (Lauritsen et al.).

Another possible reason for this difference is that urban areas, which tend to be more densely populated, have greater autism awareness. A study published in 2017 suggests that the “geographic circumstances” of living in a rural area and the decreased availability of diagnostic services result in the primary care system’s reliance on other institutions (i.e., schools) for autism-related care. The same study goes on to suggest that the comparative lack of autism awareness in rural communities, which exists on community, school, and healthcare-wide levels, results in less diagnosis (Antezana et al.). If this were true, it would be an example of how the intersection between the effect of geographic location as well as the autism diagnostic process can affect overall autism diagnosis.

Discussion

Importance of Early Diagnosis

Equitable and timely autism diagnoses are vital in ensuring the best possible care and prognosis for all autistic individuals. For children with autism, intervention is heavily focused on developing social skills and awareness, as well as on “habilitative therapies,” which are therapies focused on improving the quality of day-to-day life. Together, these therapies can help optimize the developmental path of an autistic individual. Difficult-to-correct behaviors or habits are more accessible to correct the younger the child is, making it important to diagnose the children as early as possible to give them the most time to respond to and adapt based on the therapies that are used for them (Elder et al.).

For example, a study published in 2016 supported the clinical value of Parent-mediated social communication therapy for young children with autism, or PACT. PACT was a randomized controlled trial that studied parent-mediated social communication intervention for children. The study followed up on the results of the therapy and found that it did work. However, this therapy was initially conducted on children aged two to four; this age window shows the importance of early diagnosis—children would only be eligible for this therapy if they were diagnosed at least by two years old, a relatively young age. Having robust diagnostic tools and equitable access to diagnoses would, therefore, help more children be eligible for care (Pickles et al.).

Additionally, diagnosing autism early can help families better plan for taking care of a child with autism. This includes trying different therapies and looking into other schools or programs that can better support children with autism (“Why is Early Intervention Important for Diagnosing Autism”).

Remedying the Language Barrier

One of the most important ways to ensure more equitable diagnoses for all people with autism is by fixing the language barrier issues that have been studied and proven to exist in autism testing. Specifically, as mentioned above, questionnaires are often written in one language, usually English, and then translated after the writing process is complete. While this process effectively creates a diagnostic in its original language, the meaning of the words can be lost or changed when translated. Further, specific words have different connotations in different cultures, as shown by the word spontaneous and the question about playing hide-and-seek described above.

A potential solution would be to write the original questionnaires in several languages to solve this. While the writers would work together to ensure that the tests assess the same qualities, they would write the diagnostics in different languages, ensuring that the words chosen accurately convey the intended meaning. By not making only one language the original and then translating from that model, and instead creating several “original” translations with potentially different words but more similar meanings, researchers and writers can ensure that the tests being designed are fairer and more inclusive. This increased inclusivity could affect autism diagnoses by making the diagnosis more accurate, which could, in turn, increase the speed of the diagnosis

and lower the chances of one or several incorrect diagnoses, which have detrimental impacts on the patients and their families.

Additionally, involving writers who speak different languages increases the chances of more diversity in the group of people who write any given diagnostic. This is helpful because many language barrier issues stem from cultural customs referenced in the questions (like hide-and-seek). These traditions are important to understand because autism diagnoses heavily rely on whether the patient and their caregivers understand the questions being asked in the context of their culture. More people involved from the start in the writing process means more perspectives and cultures are included in this process, which allows the diagnostics to ultimately have increased awareness for these cultural differences, resulting in the diagnostics' relatability to a particular group of people.

Healthcare Provider Awareness and Diversity

To then effectively administer these diagnostics, healthcare professionals and providers also need to have more training on how to interact with children with autism, as well as with their families. As mentioned above, there is a shortage of healthcare providers trained in diagnosing autism; this disparity is apparent considering that, in general, Black, Hispanic, and Native American people are underrepresented across several healthcare professions (Salsburg et al.). According to the American Association of Colleges of Nursing, 80% of the registered RN personnel are white (Rosseter).

These disparities need to be addressed for several reasons. First, healthcare workers can often misconstrue actual language developmental delays with not understanding the English language. A more diverse healthcare staff to administer the diagnostics could help solve this issue, as a more diverse staff who speak different languages could help distinguish between the two possibilities. Additionally, the patient-provider relationship is very important in all healthcare. It is incredibly vital in the diagnosis and treatment of autistic children specifically, as the providers need to know patients well to understand and diagnose them. Research suggests that people feel more comfortable with a healthcare provider of the same race, which emphasizes the importance of diversity—it would ensure that the patients are as comfortable as possible. This research also suggests patients are more likely to adhere to a physician's recommendations if they are of the same race, which is important for families deciding what therapies to use or further diagnoses to seek (Tanne).

It is also important that all healthcare workers, regardless of race, have a high level of training in autism diagnosis and care. Because diagnosing autism requires a trained professional, more providers must be trained in making this diagnosis. Currently, only highly trained doctors or child psychologists can diagnose autism; however, by making the training more accessible to other healthcare professionals, like nurses, a larger overall group of qualified providers would be able to diagnose autism ("Who can diagnose autism in children?"). This would also make autism diagnoses more accessible because more personnel would be capable of diagnosing in communities where it is currently more challenging to access, for example, in more rural areas.

Additionally, standardizing the guidelines for autism screening and intervention would primarily benefit healthcare workers and the families they care for. Currently, there are no international standards for autism testing, what diagnostic should be used, or what the most effective autism interventions are. The first step to solving this would be to have international guidelines for autism screening, ideally released by the World Health Organization or an equivalent institution. It would also be necessary for individual countries to release their own guidelines that consider their specific circumstances, resources, and cultural differences that could affect the autism diagnoses, such as the screening tool that makes the most sense for that culture. For these guidelines to be evidence-based, further research is needed.

Conclusion

This review has examined differences in the rate of autism diagnosis in different racial and geographical populations worldwide. It has also discussed the influence of systemic biases, internal cultural influences, and parent and caregiver influences on the diagnostic process and tools used to diagnose autism. It further looked at genetic and biological differences that can affect disparities in autism diagnoses. This review then emphasized the need for a timely diagnosis of autism and offered solutions for the language barrier and a lack of provider diversity. This review will conclude by discussing the need for further research and awareness.

Further Research

As with all topics in science, further research is needed; this is especially true with autism, as little is currently known about its causes or long-term effects. While more research needs to be done into these causes of autism generally, as well as the overall long-term effects, some new research should specifically focus on studying non-White people, as well as people not of European descent. As seen in the study that used SNPs in different subgroups, researchers have already figured out the framework for models that can be used for other groups; in that particular study, researchers focused their model on the SNPs of the CEU cohort, but if they recalibrated their diagnostic to focus on the Han Chinese SNPs, for example, it could provide a more accurate tool than the one produced by the current research.

Additionally, several methods from the study on Vitamin D study can be used—although this particular study used skin pigmentation as a proxy for Vitamin D levels, studying the actual Vitamin D levels in patients could lead to conclusive evidence about whether vitamin D levels themselves have anything to do with autism. This could then lead to discoveries about the role of immigrant status in children developing autism.

Resources and Awareness

Lastly, awareness about autism must continue to be increased around the world. Despite decreasing in the past years, stigma around autism and other neurodivergent disorders still exists. It is vital that this stigma is combated and awareness is promoted as a lack of understanding harms the person with autism, as well as their family. For a child, increased awareness can help

them fit in in varying social settings without feeling left out or excluded, which can ultimately benefit their mental health. For an adult, being more aware of the symptoms of autism can help them figure out if they should seek a diagnosis—while for adults, it may be too late for some intervention therapies, having the correct information is always beneficial. For the families and caregivers of someone who has autism, especially a child, being well-informed is incredibly important, and it allows them to make well-informed and helpful choices that will benefit the patient.

Another reason that autism awareness is important is that increased awareness leads to more resources being given to communities where access to diagnosis and care for people with autism is lacking or less than sufficient. For example, increasing autism awareness in rural communities could help more people gain accurate diagnoses. Further, raising awareness that people in rural communities do not have access to the appropriate tools could urge larger foundations and organizations to donate or send more resources to these places.

Acknowledgments

I would like to thank Dr. Carin Papendorp and Dr. Rahul Patel for their help in writing this paper.

Works Cited

- Al Shamsi, Hilal, et al. "Implications of Language Barriers for Healthcare: A Systematic Review." *Oman Medical Journal*, vol. 35, no. 2, 30 Apr. 2020, pp. 1–7, <https://doi.org/10.5001/omj.2020.40>.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed., American Psychiatric Association, 2013.
- Antezana, Ligia, et al. "Rural Trends in Diagnosis and Services for Autism Spectrum Disorder." *Frontiers in Psychology*, vol. 8, 20 Apr. 2017, www.frontiersin.org/articles/10.3389/fpsyg.2017.00590/full, <https://doi.org/10.3389/fpsyg.2017.00590>.
- Armstrong, Katrina, et al. "Racial/Ethnic Differences in Physician Distrust in the United States." *American Journal of Public Health*, vol. 97, no. 7, July 2007, pp. 1283–1289, <https://doi.org/10.2105/ajph.2005.080762>.
- Aylward, Brandon S., et al. "Racial, Ethnic, and Sociodemographic Disparities in Diagnosis of Children with Autism Spectrum Disorder." *Journal of Developmental & Behavioral Pediatrics*, vol. 42, no. 8, 10 Sept. 2021, www.ncbi.nlm.nih.gov/pmc/articles/PMC8500365/, <https://doi.org/10.1097/dbp.0000000000000996>.
- Bishop, Dorothy V. M. "Which Neurodevelopmental Disorders Get Researched and Why?" *PLoS ONE*, vol. 5, no. 11, 30 Nov. 2010, p. e15112, <https://doi.org/10.1371/journal.pone.0015112>.
- Bonham, Vence L. "Race." *Genome.gov*, 2023, www.genome.gov/genetics-glossary/Race.
- Broder-Fingert, Sarabeth, et al. "Structural Racism and Autism." *Pediatrics*, vol. 146, no. 3, 24 Aug. 2020, p. e2020015420, <https://doi.org/10.1542/peds.2020-015420>.
- Carruthers, Sophie, et al. "A Cross-Cultural Study of Autistic Traits across India, Japan and the UK." *Molecular Autism*, vol. 9, no. 1, 5 Nov. 2018, link.springer.com/article/10.1186/s13229-018-0235-3, <https://doi.org/10.1186/s13229-018-0235-3>.
- CDC. "Screening for Autism Spectrum Disorder." *Autism Spectrum Disorder (ASD)*, 8 May 2024, www.cdc.gov/autism/diagnosis/index.html.
- Constantino, John N., et al. "Timing of the Diagnosis of Autism in African American Children." *Pediatrics*, vol. 146, no. 3, 24 Aug. 2020, p. e20193629, <https://doi.org/10.1542/peds.2019-3629>.
- Dealberto, M.-J. "Prevalence of Autism according to Maternal Immigrant Status and Ethnic Origin." *Acta Psychiatrica Scandinavica*, vol. 123, no. 5, 11 Jan. 2011, pp. 339–348, <https://doi.org/10.1111/j.1600-0447.2010.01662.x>.
- Downes, Stephen M, and Lucas Matthews. "Heritability." *Stanford.edu*, 27 Feb. 2024, plato.stanford.edu/archives/spr2024/entries/heritability/. Accessed 2 Dec. 2024.
- Duello, Theresa M, et al. "Race and Genetics vs. "Race" in Genetics: A Systematic Review of the Use of African Ancestry in Genetic Studies." *Evolution, Medicine, and Public Health*, vol. 9, no. 1, 15 June 2021, www.ncbi.nlm.nih.gov/pmc/articles/PMC8604262/, <https://doi.org/10.1093/emph/eoab018>.

- “ECCM.” *W*www.eccm.org, Erie County Care Management (ECCM),
www.eccm.org/blog/why-is-early-intervention-important-for-diagnosing-autism.
- Elder, Jennifer, et al. “Clinical Impact of Early Diagnosis of Autism on the Prognosis and Parent-Child Relationships.” *Psychology Research and Behavior Management*, vol. 10, no. 10, Aug. 2017, pp. 283–292,
www.dovepress.com/clinical-impact-of-early-diagnosis-of-autism-on-the-prognosis-and-pare-peer-reviewed-fulltext-article-PRBM, <https://doi.org/10.2147/prbm.s117499>.
- Feagin, Joe, and Zinobia Bennefield. “Systemic Racism and U.S. Health Care.” *Social Science & Medicine*, vol. 103, no. 103, 19 Sept. 2013, pp. 7–14,
www.sciencedirect.com/science/article/pii/S0277953613005121,
<https://doi.org/10.1016/j.socscimed.2013.09.006>.
- Forrest, Marc P., and Peter Penzes. “Autism Genetics: Over 100 Risk Genes and Counting.” *Pediatric Neurology Briefs*, vol. 34, no. 0, 4 Dec. 2020, p. 13,
<https://doi.org/10.15844/pedneurbriefs-34-13>.
- “Gene Scoring Module.” *SFARI Gene*, gene.sfari.org/database/gene-scoring/.
- Hopkins Tanne, Janice. “Patients are more satisfied with care from doctors of same race.” *BMJ : British Medical Journal* vol. 325,7372 (2002): 1057.
- Hunt, Linda M., and Mary S. Megyesi. “The Ambiguous Meanings of the Racial/Ethnic Categories Routinely Used in Human Genetics Research.” *Social Science & Medicine*, vol. 66, no. 2, Jan. 2008, pp. 349–361, <https://doi.org/10.1016/j.socscimed.2007.08.034>. Accessed 29 June 2020.
- Jones, Desiree R, and David S Mandell. “To Address Racial Disparities in Autism Research, We Must Think Globally, Act Locally.” *Autism*, vol. 24, no. 7, 30 Sept. 2020, pp. 1587–1589, journals.sagepub.com/doi/full/10.1177/1362361320948313,
<https://doi.org/10.1177/1362361320948313>.
- Kerub, Orly, et al. “Ethnic Disparities in the Diagnosis of Autism in Southern Israel.” *Autism Research: Official Journal of the International Society for Autism Research*, vol. 14, no. 1, 1 Jan. 2021, pp. 193–201, pubmed.ncbi.nlm.nih.gov/33078561/,
<https://doi.org/10.1002/aur.2421>. Accessed 17 Sept. 2021.
- Khalid, Arshad T, et al. “Utility of Sun-Reactive Skin Typing and Melanin Index for Discerning Vitamin D Deficiency.” *Pediatric Research*, vol. 82, no. 3, 24 May 2017, pp. 444–451, www.nature.com/articles/pr2017114.pdf?origin=ppub,
<https://doi.org/10.1038/pr.2017.114>. Accessed 5 Feb. 2021.
- Kittana, Monia, et al. “The Role of Vitamin D Supplementation in Children with Autism Spectrum Disorder: A Narrative Review.” *Nutrients*, vol. 14, no. 1, 22 Dec. 2021, p. 26, <https://doi.org/10.3390/nu14010026>.
- Knerr, Sarah, et al. “Inclusion of Racial and Ethnic Minorities in Genetic Research: Advance the Spirit by Changing the Rules?” *The Journal of Law, Medicine & Ethics : A Journal of the American Society of Law, Medicine & Ethics*, vol. 39, no. 3, 2011, pp. 502–512,

- www.ncbi.nlm.nih.gov/pmc/articles/PMC4082969/,
<https://doi.org/10.1111/j.1748-720X.2011.00617.x>. Accessed 25 Feb. 2021.
- Lauritsen, Marlene B., et al. "Urbanicity and Autism Spectrum Disorders." *Journal of Autism and Developmental Disorders*, vol. 44, no. 2, 1 Feb. 2014, pp. 394–404,
www.ncbi.nlm.nih.gov/pmc/articles/PMC4511947/,
<https://doi.org/10.1007/s10803-013-1875-y>.
- Mandell, David S., et al. "Disparities in Diagnoses Received prior to a Diagnosis of Autism Spectrum Disorder." *Journal of Autism and Developmental Disorders*, vol. 37, no. 9, 8 Dec. 2006, pp. 1795–1802, www.ncbi.nlm.nih.gov/pmc/articles/PMC2861330/,
<https://doi.org/10.1007/s10803-006-0314-8>.
- . "Racial/Ethnic Disparities in the Identification of Children with Autism Spectrum Disorders." *American Journal of Public Health*, vol. 99, no. 3, Mar. 2009, pp. 493–498,
www.ncbi.nlm.nih.gov/pmc/articles/PMC2661453/,
<https://doi.org/10.2105/ajph.2007.131243>.
- Mazurek, Micah O., et al. "ECHO Autism STAT: Accelerating Early Access to Autism Diagnosis." *Journal of Autism and Developmental Disorders*, vol. 49, no. 1, 24 July 2018, pp. 127–137, <https://doi.org/10.1007/s10803-018-3696-5>.
- McCleary, Lynn, and Mark Sanford. "Parental Expressed Emotion in Depressed Adolescents: Prediction of Clinical Course and Relationship to Comorbid Disorders and Social Functioning." *Journal of Child Psychology and Psychiatry*, vol. 43, no. 5, July 2002, pp. 587–595, <https://doi.org/10.1111/1469-7610.00048>. Accessed 30 Nov. 2021.
- McNally Keehn, Rebecca, et al. "A Statewide Tiered System for Screening and Diagnosis of Autism Spectrum Disorder." *Pediatrics*, vol. 146, no. 2, 6 July 2020, p. e20193876,
<https://doi.org/10.1542/peds.2019-3876>.
- Nielsen-Bohlman, Lynn. *Health Literacy : A Prescription to End Confusion*. Washington, D.C., National Academies Press, 2004.
- Office of the Surgeon General (US), et al. *Mental Health: Culture, Race, and Ethnicity: A Supplement to Mental Health: A Report of the Surgeon General*. National Library of Medicine, 2001, www.ncbi.nlm.nih.gov/books/NBK44249/.
- Pattillo, Mia, et al. "Racism against Racialized Migrants in Healthcare in Europe: A Scoping Review." *International Journal for Equity in Health*, vol. 22, no. 1, 29 Sept. 2023,
<https://doi.org/10.1186/s12939-023-02014-1>.
- Pickles, Andrew, et al. "Parent-Mediated Social Communication Therapy for Young Children with Autism (PACT): Long-Term Follow-up of a Randomised Controlled Trial." *The Lancet*, vol. 388, no. 10059, Nov. 2016, pp. 2501–2509,
[www.thelancet.com/journals/lancet/article/PIIS0140-6736\(16\)31229-6/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(16)31229-6/fulltext),
[https://doi.org/10.1016/s0140-6736\(16\)31229-6](https://doi.org/10.1016/s0140-6736(16)31229-6).
- "Recommendation: Autism Spectrum Disorder in Young Children: Screening | United States Preventive Services Taskforce." *Www.uspreventiveservicestaskforce.org*, 2016,

- www.uspreventiveservicestaskforce.org/uspstf/recommendation/autism-spectrum-disorder-in-young-children-screening.
- Rosseter, Robert. "Nursing Workforce Fact Sheet." *Wwww.aacnnursing.org*, American Association of Colleges of Nursing, Apr. 2024, www.aacnnursing.org/news-data/fact-sheets/nursing-workforce-fact-sheet.
- Salsberg, Edward, et al. "Estimation and Comparison of Current and Future Racial/Ethnic Representation in the US Health Care Workforce." *JAMA Network Open*, vol. 4, no. 3, 31 Mar. 2021, <https://doi.org/10.1001/jamanetworkopen.2021.3789>.
- Sandin, Sven, et al. "The Heritability of Autism Spectrum Disorder." *JAMA*, vol. 318, no. 12, 26 Sept. 2017, p. 1182, jamanetwork.com/journals/jama/fullarticle/2654804, <https://doi.org/10.1001/jama.2017.12141>.
- Skafidas, Stan, et al. "Predicting the Diagnosis of Autism Spectrum Disorder Using Gene Pathway Analysis." *Molecular Psychiatry*, vol. 19, no. 4, 11 Sept. 2012, pp. 504–510, <https://doi.org/10.1038/mp.2012.126>. Accessed 28 Apr. 2020.
- St. Amant, Helaine G., et al. "Language Barriers Impact Access to Services for Children with Autism Spectrum Disorders." *Journal of Autism and Developmental Disorders*, vol. 48, no. 2, 7 Oct. 2017, pp. 333–340, <https://doi.org/10.1007/s10803-017-3330-y>.
- Stahmer, Aubyn C., et al. "Caregiver Voices: Cross-Cultural Input on Improving Access to Autism Services." *Journal of Racial and Ethnic Health Disparities*, vol. 6, no. 4, 11 Mar. 2019, pp. 752–773, <https://doi.org/10.1007/s40615-019-00575-y>.
- Tabassum, Rubina, et al. "Geographical, Environmental and Pathophysiological Influences on the Human Blood Transcriptome." *Current Genetic Medicine Reports*, vol. 1, no. 4, 1 Oct. 2013, pp. 203–211, <https://doi.org/10.1007/s40142-013-0028-0>. Accessed 16 July 2021.
- Templeton, Alan R. "Biological Races in Humans." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 44, no. 3, Sept. 2013, pp. 262–271, www.ncbi.nlm.nih.gov/pmc/articles/PMC3737365/, <https://doi.org/10.1016/j.shpsc.2013.04.010>.
- Tick, Beata, et al. "Heritability of Autism Spectrum Disorders: A Meta-Analysis of Twin Studies." *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, vol. 57, no. 5, 2016, pp. 585–95, pubmed.ncbi.nlm.nih.gov/26709141/, <https://doi.org/10.1111/jcpp.12499>.
- "When Do Children Usually Show Symptoms of Autism?" *Http://Wwww.nichd.nih.gov/*, Eunice Kennedy Shriver National Institute of Child Health and Human Development, 31 Jan. 2017, www.nichd.nih.gov/health/topics/autism/conditioninfo/symptoms-appear.
- "Who Can Diagnose Autism in Children? – Diagnosing Autism - the Place for Children with Autism." *The Place for Children with Autism*, 24 Oct. 2024, theplaceforchildrenwithautism.com/diagnosing-autism/who-can-diagnose-autism. Accessed 2 Dec. 2024.
- Zeidan, Jinan, et al. "Global Prevalence of Autism: A Systematic Review Update." *Autism Research*, vol. 15, no. 5, 2022, pp. 778–790, onlinelibrary.wiley.com/doi/10.1002/aur.2696, <https://doi.org/10.1002/aur.2696>.

The Ugly Truth Behind the Largest Enterprises: How Are Small Businesses (SMEs) Being Impacted? By Samanvi Hublikar

Abstract

The purpose of this study is to examine the effects that large businesses have on the small business economy. In this paper, the following topics are discussed: How oligopolies affect small businesses, the lack of competitive advantage for small businesses, killer acquisitions, and idea theft. To create this literature review, sources that were peer-reviewed, credible, and published between 2020-2024 are used. Search phrases such as “small business economy,” “large corporations,” “large businesses exploiting small businesses,” were used. The conclusion includes unethical operations that large businesses conduct, unbalanced playing fields, and future implications such as establishing solvency.

Introduction

Imagine a world where you, as a customer, can't control the market and thus can not efficiently access your needs. Due to the accelerating power of large corporations, this is the current status of the economy in the United States. The diversity of businesses due to upcoming small businesses has always led to a consumer-controlled market. However, during the economic crisis amid the coronavirus pandemic, many large companies have grown rapidly while their small business competitors have been facing an apocalypse (Goolsbee, 2020).

Small to mid-size enterprises (SME; also known as small businesses) are critical for the economic ecosystem. SMEs represent companies with fewer than 500 employees, and around 50% of small businesses in the US have between 1 to 4 employees (Leppert, 2024). Small businesses contribute to 44% of the US economy (U.S. Chamber of Commerce, May 2024) and represent almost all of the businesses in the US as approximately 99.9% of businesses are small businesses (Lucy, Hoover, Furgerson, 2024). Therefore, it is clear that small businesses are the fuel to significantly boost the US economy.

Despite the innovation that small businesses bring to the US economy, small businesses have failed to survive past the five-year mark (Goodarzi, 2024; U.S. Chamber of Commerce, May 2024). With large businesses' economies of scale, small businesses face detrimental threats such as idea theft, exploitation, and killer acquisitions, increasing monopolies that limit consumer choice, which critically threaten innovation. Large businesses significantly affect small businesses poorly due to the gaping imbalances of power, affecting not only these small corporations but also the average customer citizen.

The goal of this paper is to examine the impact of large corporations' economic power on the small business economy in the United States. This analysis will focus on the period after the coronavirus pandemic to present the most up-to-date picture of the economic ecosystem. Based on the existing literature, case studies, and systematic developments, we predict that the rise of large corporations negatively affects the small business economy (SME) for two main reasons:

1. First, small businesses don't have a competitive advantage in an ecosystem where large corporations dominate the market and create a monopoly.
2. Second, although small businesses are known for driving innovation, large companies often take the creativity credit because they have the financial and logistical resources to translate innovative ideas into reality.

Methodology

In this literature review, a systematic approach was employed to synthesize relevant literature on the effects of large businesses on the small business economy.

The literature search was conducted using Google Scholar. The search terms and keywords used included “small business economy,” “large corporations,” “large businesses exploiting small businesses,” “the power of large businesses against small businesses,” “economic ecosystem,” “company acquisition,” “what is the current US economy like for small business?”

The following inclusion and exclusion criteria were used to select the studies for review: Inclusion criteria: Studies published between 2020-2024. Peer-reviewed articles and conference papers. Literature focused on monopolies, business practices, competitive advantages, etc.

Exclusion criteria: Non-peer-reviewed sources, such as editorials, blogs, and opinion pieces.

Studies not available in English. Studies that did not focus directly on the direct effect of large businesses on small businesses.

A multi-step process was followed to ensure a comprehensive and unbiased selection of literature. First, the titles and abstracts of all retrieved articles were screened. After an initial screening, full-text articles were assessed for relevance to the research question. Duplicates were removed, and a final list of 11 studies was selected for review.

The selected studies were analyzed using a thematic synthesis approach. Studies were grouped based on the topic of the current paper, and key patterns, trends, and themes were identified. Discrepancies between studies were highlighted and discussed where relevant. This review is limited by the exclusion of non-English language papers, studies older than 2020, and by the fact that only peer-reviewed publications were considered, potentially excluding valuable insights from gray literature.

Results

Market Share Dominance

Successful large enterprises seem like something that you have to immediately respect. However, the power they hold can create an unbalanced playing field with the potential for power abuse. Small businesses are negatively and constantly affected by such tactics that large businesses use to keep their power. Therefore, large companies are harming the small business economy due to the lack of a competitive advantage that the small businesses hold.

In an ecosystem where large corporations dominate the market and create a monopoly, it is challenging for SMEs to compete. Therefore, many small businesses find it hard to answer the question: how can we become profitable when there is an already established oligopoly (when the market is controlled by a small number of businesses, limiting competition)? With significant market share and killer acquisitions, large businesses hurt small businesses and diminish consumer choice.

Large businesses dominate through market share. Tech giants such as Amazon and Apple have already acquired a significant amount of market share. The growing power of such large companies has led to difficult grounds for start-ups to compete (Lee, 2020). In return, start-ups find it difficult to stay alive and eventually get disbursed, contributing to another stepping stone to an undiversified landscape. Popular firms also discourage upcoming businesses, allowing the popular firms to continue to claim all the power. The power imbalance between large businesses and SMEs has also led to a dependence on these established businesses. For example, many small businesses depend on Amazon to sell their goods, allowing Amazon to extract money when a sale is made on its platform (Velazquez, 2024).

Killer Acquisitions

Larger companies also participate in killer acquisitions, strengthening the large business's power. These businesses easily purchase innovative competitors to remove any threats (Goolsbee, 2020). An example of a killer acquisition can be seen through Facebook's takeover of WhatsApp, an instant messaging app. Facebook's motivation for taking over WhatsApp was to eliminate a competitor and protect its power (Stek, 2023). Additionally, in many instances, small businesses that are acquired are even destroyed, absorbed, or vanished (Patel, 2021). Thus, not only do killer acquisitions take over a small business, some acquisitions also destroy innovation.

From these power imbalances and tactics, plenty of small businesses are destroyed, and many people are even discouraged from starting a business. This ever-growing increase in market concentration is leading to more monopolies. Not only do monopolies hurt SMEs, but they also hurt consumers through inefficiencies, reductions in innovation, and high prices (Kelly, 2024).

For all these reasons, it is clear that an unbalanced competitive landscape due to the sheer amount of power that large corporations hold leads to an undiversified landscape, hurting small business owners and consumers.

Idea Theft

The average consumer might admire new innovative products that large corporations produce. Well, what if I tell you that the world's most famous cookie, Oreo, was stolen by a company called Hydrox? Unfortunately, this case isn't the only one as many large corporations engage in this unethical practice. Although small businesses are known for driving innovation, large companies often take the creativity credit because they have the financial and logistical resources to translate innovative ideas into reality.

When large companies discover innovative ideas from SMEs, they often commit idea theft, leaving the SMEs with their work stolen. Unfortunately, small businesses are not able to compete with these large businesses due to the large businesses' economies of scale. Large businesses are then able to sell the stolen item at a cheaper price, potentially driving the SME out of business.

An example of idea theft can be seen in a notorious company, Shein. At one point, Shein was reported to have stolen art from a small artist six times. If that artist tried to talk to Shein about the copyright claims, Shein would simply pay some money for an unlimited license and delete the products (Liu, 2022). Unfortunately, stealing ideas is ingrained into plenty of fast fashion systems as a complaint displays that fast fashion companies combine a data mining system with surveillance of its competitors (Deppen, 2024). To add on, other big fast fashion companies such as Zara and H&M are constantly involved in hundreds of lawsuits annually about the similarity of their works to others.

Large companies are able to copy products from other businesses through multiple ways. These bigger companies are able to poach employees from small businesses to get information, get access to confidential information, use tactics such as predatory pricing to make small businesses more vulnerable, and exploit loopholes (FasterCapital, 2024). To add on, it is extremely easy for large businesses to be able to participate in these unethical tactics and strategies due to the amount of resources and capital they have accumulated. Thus, even if the SMEs are able to produce innovative ideas that would potentially allow them to break into the market and create a name for themselves, large businesses can stop this process by engaging in unethical behavior, driving small businesses to the ground.

Discussion

Large enterprises' power hurts small businesses from being profitable and sustainable. Not only do the power imbalances hurt SMEs, but unethical practices that large businesses engage such as killer acquisitions and idea theft make it that much harder for SMEs to compete. This stifles innovation as smaller companies with creative ideas are often crushed before they can fully enter the market. Moreover, these practices contribute to monopolistic environments, leaving consumers with fewer choices and higher prices.

The implications of this paper include encouraging large businesses to start operating ethically. Instead of performing idea theft and killer acquisitions, larger businesses should allow for a fair playing field. By fostering a competitive environment, they could also benefit from healthy competition, which drives innovation and market growth. Ethical behavior would also improve the overall perception of large corporations, allowing trust and long-term sustainability.

Current limitations of this literature review are that outside variables that affect small businesses such as investments and starter funds are not discussed. Large businesses do play a significant role in small business destruction, however, other variables do play roles as well. Factors like government regulations, access to technology, and consumer behavior also need to be considered to understand the full scope of the challenges faced by SMEs. Without accounting

for these external factors, the analysis may overlook key components that contribute to the struggles of small businesses.

Future research could further explore understanding killer acquisitions and other unethical issues. This paper covers many topics and allows for a bigger picture understanding. However, it is also important for readers to have a detailed understanding of unethical tactics businesses use. Future research could also benefit from case studies of specific SMEs affected by these practices, providing a more concrete understanding of the real-world consequences.

In conclusion, the effects go far from only hurting small business owners. Destroying small businesses leads to an undiversified economic landscape that creates monopolies and leads to a lack of innovation and the absence of consumer choice. Overall, it is impossible and unwise to have large businesses give up their market share. However, there is an unfair playing ground when it comes to small businesses trying their best to thrive. In the end, large businesses once started as SMEs, and they should understand the struggles and make sure to perform ethically to operate in good faith and to create a slightly more equal playing field in a highly unbalanced one.

Works Cited

- Deppen, Laurel. "Shein Has Copyright Infringement 'baked into' Its Business Model: Class Action Lawsuit." *Fashion Dive*, 10 Apr. 2024, www.fashiondive.com/news/shein-class-action-lawsuit-copyright/712798/.
- Ferguson, Stephanie, et al. "See the Data behind America's Small Businesses." *See the Data behind America's Small Businesses*. | U.S. Chamber of Commerce, U.S. Chamber of Commerce, 5 Sept. 2024, www.uschamber.com/small-business/small-business-data-center.
- Goolsbee, Austan. "Big Companies Are Starting to Swallow the World." *The New York Times*, The New York Times, 30 Sept. 2020, www.nytimes.com/2020/09/30/business/big-companies-are-starting-to-swallow-the-world.html.
- Goodarzi, Sasan. "Small Businesses Are Cutting Jobs. It's a Warning Sign for the US Economy." *USA Today*, Gannett Satellite Information Network, 14 Mar. 2024, www.usatoday.com/story/opinion/2024/03/14/economy-small-businesses-interest-rates-credit/72955958007/.
- Lee, Don, "Big Companies Are Swallowing up Smaller Ones. Why That's Not Good for the Economy." *Los Angeles Times*, Los Angeles Times, 27 Sept. 2020, www.latimes.com/politics/story/2020-09-27/big-companies-increasingly-swallow-competitors-could-hurt-economy.
- Marieke. "Breaking down Killer Acquisitions: The Rise of below-Threshold Merger Scrutiny." *Stek Lawyers*, 1 May 2024, stek.com/en/2023/05/16/breaking-down-killer-acquisitions-the-rise-of-below-threshold-merger-scrutiny/.
- Patel, Kison, "Do Big Companies Destroy Small Ones in the M&A Process?" *FEI*, www.financialexecutives.org/FEI-Daily/May-2021/Do-Big-Companies-Destroy-Small-Ones-in-the-M-A-Pro.aspx.
- Team, The Investopedia. "Are Monopolies Always Bad?" *Investopedia*, Investopedia, www.investopedia.com/ask/answers/032615/are-monopolies-always-bad.asp#:~:text=A%20monopoly%20exists%20when%20one,of%20innovation%2C%20and%20higher%20prices.
- Velazquez, Rep. Nydia M. (D-N.Y.), opinion contributor. "Corporate Concentration Threatens America's Small Business Boom." *The Hill*, The Hill, 13 Mar. 2024, thehill.com/opinion/finance/4527961-corporate-concentration-threatens-americas-small-business-boom/.
- FasterCapital, "Sneaky Ways Big Companies Are Stealing Small Businesses Ideas", FasterCapital, June 2024, fastercapital.com/content/Sneaky-ways-big-companies-are-stealing-small-businesses-ideas.html.
- Liu, Jiaye, "Research on the Business Strategy and Deficiency of the Fast Fashion Industry to Enhance Development - a Case Study of Shein," King's Business School, <https://www.google.com/url?q=https://www.atlantis-press.com/article/125983743.pdf&sa=D&source=docs&ust=1726445331437892&usq=AOvVaw04ByoBc4NfNmGVS5Cv2wuu>

Optimal Tilt Angle For Solar Collectors In Telangana, India By Varun Ponnala

Abstract

As the world faces problems of depleting natural resources and calls for sustainable energy alternatives, solar energy has become an important viable source. Effective use of solar energy requires determining the maximum amount of solar radiation the collectors can receive by obtaining the optimal tilt angle of solar collectors. This paper aims to find the optimal tilt angle of solar collectors in Shamshabad, Telangana, India.

Experimental readings were collected using the Lux Light Meter app to measure the incoming solar radiation at various tilt angles on a farm in Shamshabad. Measurements were taken over a span of one month to capture variations in the solar radiation received on a solar collector. The methodology allowed for systematic evaluation of how the tilt angle influences solar radiation capture.

The obtained results make it possible for the optimization of the incoming solar radiation received by solar collectors in Telangana, taking into account practical insights to optimize solar energy in the region. Comparing the theoretically calculated tilt angle with the experimentally obtained ones showed rather a good similarity to establishing a conclusion for the period studied. There were some small discrepancies, which emphasized that adjustments should be made according to the local environmental conditions. Such results open practical avenues for exploitation in designing solar installations and provoke further investigation into seasonal adjustments, so that the efficiency may be further increased.

Keywords Optimal tilt angle; Solar collectors; Solar radiation; Telangana

Introduction

The world is facing an enormous energy crisis because of the rapid depletion of non-renewable resources like coal, oil, and natural gas. In the light of this event, renewable energy sources are turning out to be the most critical for meeting global energy demands sustainably. India has a good geographic condition: abundant sunlight all year round; it is one of the most promising regions for solar energy development. The place, Shamshabad, 17.2595° N latitude and 78.3967° E longitude, in Telangana, India, is one of these locations with potential for high solar radiation, making solar energy a viable choice. Maximizing the accumulation of solar energy available using the solar collector system requires finding the optimal tilt angle.

Duffie and Beckman [2] provide important theoretical frameworks for solar energy studies, such as diagrams to understand derived equations like the $\cos \theta$ relation, which helps in deriving the formula for the optimal tilt angle for a solar collector. Handoyo et al. [4] investigated the optimal tilt angle for solar collectors over Indonesia and emphasized that maximizing the cosine of the incident angle ensures maximum solar radiation absorption. Their results validated with empirical values yield a variation of optimal tilt angles based on seasonal shifts. Safdarian et al. [5] investigated the best tilt and azimuth angles of solar collectors

across main cities in Iran using a Gravitational Search Algorithm (GSA). Their results highlighted the importance of dynamic adjustments to tilt angles to maximize solar radiation, especially on days with clear weather. Tang et al. [6] developed a simplified mathematical model to estimate optimal tilt angles for fixed solar collectors in China, mapping annual tilt values for 152 locations. Their results confirmed the reliability of their method with respect to the location where limited diffuse radiation data are available. Emanuele Calabrò [1] examined the relationship between latitude and optimal tilt angles for photovoltaic panels in the north-tropical latitudes. He recommended a seasonal angle adjustment twice a year to account for the considerable seasonal variations he found for his study site in Italy. Gunerhan et al. [3] concluded Turkey-based experimental research by suggesting that monthly changes in tilt angle are necessary for maximizing the solar collector efficiency especially for building applications. Together, these studies give an emphasis on local experimentation of solar energy and adjustment of tilt angles in maximizing incoming solar radiation received on a solar collector.

In this paper, the solar radiation on different tilt angles is measured with the Lux Light Meter Pro application, version 036.2024.05.05 (available on the Play Store). To ensure accuracy in the readings, the tilt angle of the phone is adjusted with a protractor during the reading and a cardboard is placed under the phone for stability as in Figure 1.



Fig 1. Experimental Setup

The measurements were taken over a period of one month from 10:00 AM to 1:00 PM on four days chosen at intervals of roughly 15 days, namely September 22nd, 2024, October 7th, 2024, October 22nd, 2024, and November 16th, 2024. The incoming solar radiation was measured by standing beside pre-installed solar panels at a farm in Shamshabad, in actual conditions to imitate real solar collector setups. This systematic approach ensured that measurements were practical and directly applicable to the optimization of the capture of solar energy within the region.

Nomenclature

θ	Angle of incidence ($^{\circ}$)
β	Tilt angle of collector with respect to horizontal ($^{\circ}$)
δ	Declination of the sun ($^{\circ}$)
ϕ	Latitude ($^{\circ}$)
γ	Azimuth angle ($^{\circ}$)
ω	Hour angle ($^{\circ}$)
θ_z	Zenith angle ($^{\circ}$)
G_b	Beam radiation intensity on a surface (W/m^2)
G_{bn}	Beam radiation intensity at normal incidence (W/m^2)
G_{bt}	Beam radiation intensity on a tilted surface (W/m^2)

Theoretical Development

The relation between the angle of incidence, tilt angle, zenith angle and the solar beam radiation on a horizontal and tilted surface is depicted in Figure 2 [2]. The angle of incidence can be described as the angle formed between the solar rays and the normal to the surface. This angle is highly important in determining the amount of beam radiation (G_{bt}) that reaches an inclined surface relative to the amount of beam radiation that is perpendicular to the sun (G_{bn}).

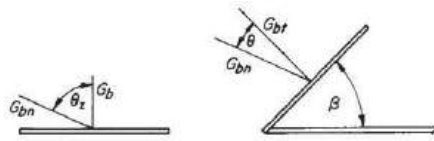


Fig 2. Beam radiation on horizontal and tilted surfaces.

Forming a trigonometric relationship, we get the beam radiation that falls on the tilted surface as $G_{bt} = G_{bn} \cos \theta$ where $\cos \theta$ determines the projection of G_{bn} onto a slanted plane.

$$\cos \theta = \sin \delta \sin \phi \cos \beta - \sin \delta \cos \phi \sin \beta \cos \gamma + \cos \delta \cos \phi \cos \beta \cos \omega + \cos \delta \sin \phi \sin \beta \cos \gamma \cos \omega + \cos \delta \sin \beta \sin \gamma \sin \omega \quad (1)$$

Equation (1) relates to the angle of incidence, with solar radiation striking an inclined surface. There are several geometric relationships of angles involved in (1). These factors describe the sun's position with respect to a solar collector and determine how the tilt, orientation and solar position affect the incoming solar radiation that interacts with the surface.

To measure the amount of solar radiation, the Light Lux Meter Pro application was used. It measures light intensity in lux. To ensure that the readings taken are accurate, the smartphone was placed on a piece of cardboard (acting as a solar panel) flat on the surface and slightly inclined by a protractor to the following angles: 0°, 10°, 20°, 30°, 40°, 50°, 60°, 70°, 80°, and 90°. The values of lux obtained through the application were converted into estimated irradiance values using the following conversion formula: $1 \text{ lux} \approx 0.0083 \text{ W/m}^2$. The theoretical values of $\cos \theta$ were calculated using (1), thus computing declination, latitude, tilt of the surface, azimuth, and hour angles. As $\cos \theta$ directly affects the intensity of beam radiation hitting the inclined surface, a comparative analysis pointed out the relationship between the $\cos \theta$ values and solar radiation. The readings were taken on September 22nd at 10AM, October 7th at 1PM and November 16th at 11:45 AM of 2024.

For (1), the declination angle (δ) can be calculated as

$$\delta = 23.45^\circ \times \sin\left(\frac{360}{365} \times (n - 81)\right) \quad (2)$$

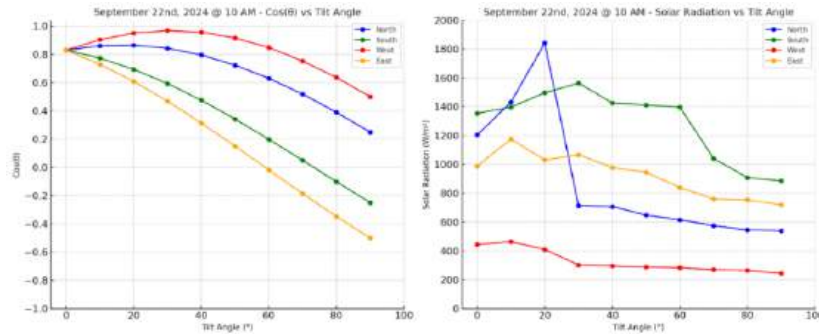
where n is the nth day of the year.

The hour angle (ω) can be calculated by

$$\omega = 15^\circ \times (\text{Solar Time} - 12) \quad (3)$$

where the solar time represents the time of the day based on the position of the sun.

The azimuth angle (γ) for different surface orientations is 0° for North, 90° for East, 180° for South and 270° for West.



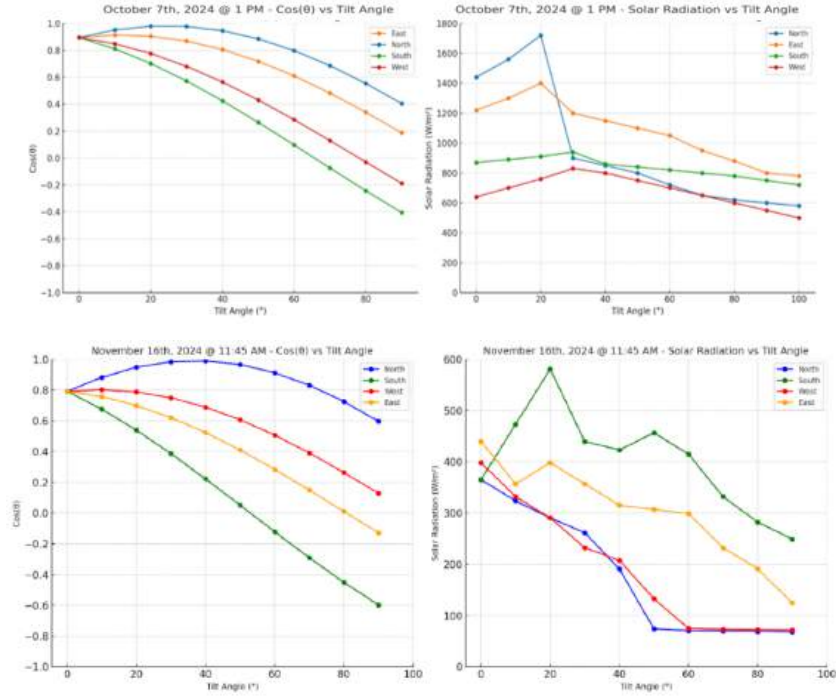


Fig 3. $\cos \theta$ calculated from (1) and the measured incoming solar radiation received on several tilt angles for the chosen days

Figure 3 reflects that when $\cos \theta$ reduces, the measured incoming solar radiation also reduces. This indicates that $\cos \theta$ is proportional to the solar radiation observed. Thus, maximizing $\cos \theta$ on a surface can maximize the solar radiation received on that surface.

Calculating the optimal tilt angle

Since maximizing $\cos \theta$ can maximize the solar radiation received, the tilt angle (β) that maximizes $\cos \theta$ would be the optimal tilt angle. The optimal tilt angle value can be obtained from making $\left(\frac{d\cos \theta}{d\beta}\right) = 0$ [4]. After deriving an equation for β , we get

$$\tan \beta = \frac{-\sin \delta \cos \phi \cos \gamma + \cos \delta \sin \phi \cos \gamma \cos \omega + \cos \delta \sin \gamma \sin \omega}{\sin \delta \sin \phi + \cos \delta \cos \phi \cos \omega} \quad (4)$$

This β value will make $\cos \theta$ maximum when $\frac{d^2 \cos \theta}{d\beta^2} < 0$.

Calculating the second derivative we get

$$\frac{d^2 \cos \theta}{d\beta^2} = -A \cos \beta - B \sin \beta \quad (5)$$

In (5),

$$A = (\sin \delta \sin \phi + \cos \delta \cos \phi \cos \omega) ,$$

$$B = (-\sin \delta \cos \phi \cos \gamma + \cos \delta \sin \phi \cos \gamma \cos \omega + \cos \delta \sin \gamma \sin \omega)$$

The tilt angle (β) can be calculated through (4), only requiring the day of the year (declination angle), location (latitude), orientation of the surface (surface azimuth angle), and the time of the day (hour angle). To check if the second derivative was negative, the second derivative values were computed using an excel spreadsheet for different days, timings, and different surface orientations. From the values computed, it was found that all the second derivative values were negative, making β the optimal tilt angle as it maximizes the value of $\cos \theta$.

In Figures 4 and 5, another set of readings were taken at different timings (10AM, 11AM, 12PM and 1PM) on 3 days (September 22nd, October 7th and October 22nd of 2024) to experimentally obtain an optimal tilt angle.

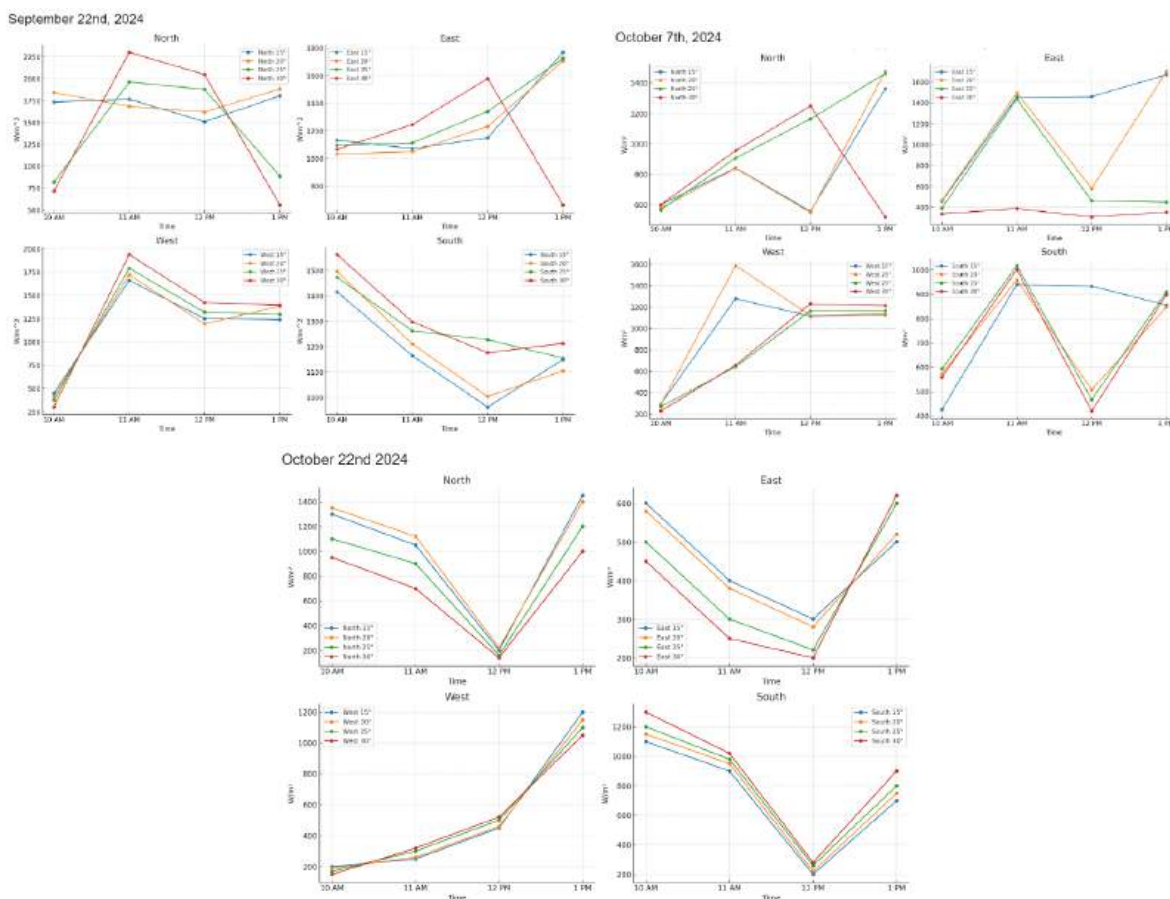


Fig 4. Incoming solar radiation measured on three chosen days

Analyzing the solar radiation data and the various tilt angles used for the dates, it is found that a collector can be mounted at an optimal tilt angle varying between 18° to 28° to the North

to maximize solar radiation collection. An alternate method involves a dynamic adjustment of the orientation and tilt method where a solar collector should face towards the East in the morning and West in the afternoon (after 12PM) with an optimal tilt angle ranging from 15° to 30° for both orientations. Both the calculations from (4) and the experimentally obtained optimal tilt angle support these results.

Discussion

An analysis of the solar radiation data shows that a solar collector can be positioned at a fixed optimal tilt angle to maximize solar energy collection, or dynamically adjusted to face different directions during the day to optimize the incoming solar radiation better. There is some scope for improvement in the data collection process. Through gathering data from the Lux Light Meter Pro application and measuring the amount of light intensity in lux. Converting the lux into the solar radiation unit (W/m^2), valuable insights were provided. However, if a more accurate tool like a pyranometer was used, it could have been used to directly measure the incoming solar radiation in W/m^2 . In addition, the varying weather conditions during the dates when the data were accumulated, may have created discrepancies in data from the clouds covering the sun or dust in the atmosphere. Furthermore, the data collected was over a span of 1 month. This doesn't take into account the seasonal variations and how it could be required to take a different optimization approach for each season throughout the year. In the future, more research could be carried out by using more accurate equipment and taking into consideration the seasonal and daily changes in weather to improve accuracy in data collection.

Conclusion

Through comparing the solar radiation data along with the $\cos \theta$ value, this study was able to prove that $\cos \theta$ is proportional to the solar radiation. The calculated values in (4) compare well with the experimental values obtained for determining the optimum tilt angle of a solar collector in Telangana, India. The solar radiation data at various dates and tilt angles suggest that the fixed arrangement where the collector is placed towards North receives the highest solar radiation, which falls within the optimal tilt range of 18° to 28° . Alternatively, in the morning, the collector can face east and face west in the afternoon (after 12PM). For both orientations, the optimal tilt angles vary from 15° to 30° .

Dynamic adjustments are an improvement in the efficiency of optimizing the incoming solar radiation, however, it is not so practical due to the frequent changes required. This may not be feasible due to the additional time and effort necessary to achieve this orientation. The North-facing approach is better in terms of its simplicity in its implementation because it does not demand frequent changes and still offers great efficiency in gathering incoming solar radiation.

The optimal tilt angle changes throughout the year to different weather conditions. From this, one question arises: Will future developments make adjustable solar panels the norm, or will fixed systems that are improved for local conditions be the way forward?

Works Cited

- Calabro, E. (2009). Determining optimum tilt angles of photovoltaic panels at typical north-tropical latitudes. *Journal of Renewable and Sustainable Energy*, 1, 033104.
doi: 10.1063/1.3148272.
- Duffie, J.A., & Beckman, W.A. (2020). *Solar Engineering of Thermal Processes*. John Wiley & Sons, Inc.
- Gunerhan, H. & Hepbasli, A. (2007). Determination of the optimum tilt angle of solar collectors for building applications. *Building and Environment*, 42(2), 779-783.
doi: 10.1016/j.buildenv.2005.09.012
- Handoyo, E. A., Ichsani, D., & Prabowo. (2013). The Optimal Tilt Angle of a Solar Collector. *Energy Procedia*, 32, 166–175.
doi: 10.1016/j.egypro.2013.05.022.
- Safdarian, F., & Nazari M. E. (2015). Optimal tilt angle and orientation for solar collectors in Iran. *2015 IEEE 10th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, Guarda, Portugal, 494-500.
doi: 10.1109/DEMPED.2015.7303735
- Tang, R. & Wu, T. (2004). Optimal tilt-angles for solar collectors used in China. *Applied Energy*, 79, 239-248.
doi:10.1016/j.apenergy.2004.01.003.

Choking Civic Spaces: Digital Censorship and Its Effects on Indonesian Civil Society Organizations

By Jayden Christian Budiman

Abstract

The problematic introduction and implementation of Indonesia's laws on digital censorship marks the nation's backslide into authoritarianism as civil society organizations (CSOs) are being stripped of their power to hold the government accountable for abuses of human rights. Existing research has primarily focused on the global rise of suppression of individuals and their ability to pursue free expression. However, in the specific geographical context of Indonesia, there is a severe lack of literature regarding its impacts on CSOs, the most important catalysts of social change and government oversight. Through the use of qualitative data, including analysis of legal documents, case studies, and interviews with CSO leaders, this paper provides a nuanced understanding of (a) how CSOs leverage the internet in Indonesia's increasingly monopolized media landscape and (b) how censorship laws affect their advocacy efforts. The findings reveal that digital censorship has been weaponized against government critics, inducing a chilling effect on CSOs and stripping Indonesians of a vital source of truthful journalism. Simultaneously, these laws have failed in their stated purpose as the prominence of harmful online rhetoric against marginalized communities continues to grow in Indonesia. Such a phenomenon highlights the need for inclusive and transparent reforms to censorship policies to safeguard democratic values in Indonesia's digital age.

Introduction

Digital censorship commonly refers to attempts to limit access to internet content to control information that can be viewed or published online. Although the United Nations Human Rights Council has condemned any form of digital censorship as a heinous violation of an individual's right to freedom of expression, this has not stopped numerous authoritarian states across the world from clamping down on online discourse that challenges their rule (Bischoff 2023). Worryingly, despite the mainstream belief that society is liberalizing, digital censorship has been on the rise. 2023 marked the thirteenth consecutive year that saw a decline in worldwide digital freedom, with a record high 55 countries now have laws that criminalize online expression (Henshall 2023). Ironically, this was caused by recent technological advances which have hurt instead of helping freedom on the internet. Indonesia, with its dismal internet freedom score of 49 (a score of 0 being the least free and one of 100 being the freest), is of no exception (Freedom House 2024).

As one of the world's most populous and rapidly developing nations, Indonesia boasts a staggering internet user base of 240 million out of a total population of 275 million as of 2022, and this number is only projected to grow in the coming years ("Indonesia: Number of Internet Users 2028 | Statista" 2023). The ubiquitous nature of internet use has made social media a fundamental channel of information. Consequently, civil society organizations (CSOs) in Indonesia have constantly relied on the digital world as its primary means of communication and

method to raise awareness on a multitude of human rights issues such as labor rights, women's empowerment, and the protection of minorities. This is especially important in a developing country such as Indonesia which has a mediocre fundamental rights score of 0.5 (a score of 0 indicating a very weak adherence to fundamental rights and a score of 1 indicating strong respect for fundamental rights) (WJP Rule of Law Index 2024). CSOs acting as a source of truthful and impartial information play a crucial role in upholding democratic practices in Indonesia because they keep the government accountable for their actions. These CSOs include the likes of KontraS, whose journalists monitor instances of state violence against marginalized communities, as well as Amnesty International Indonesia, which publicizes data on human rights concerns in Indonesia. The widespread use of the internet by CSOs as a tool of news has elicited swift and targeted retaliations from the government, advancing digital censorship in the country.

Under the Jokowi Administration (2014-2024), the level of digital censorship in Indonesia has grown increasingly stringent to the detriment of society, most frequently manifesting in unnecessarily broad and vague laws introduced under the auspices of preserve national security, namely, the Electronic and Information Law (ITE Law) and Ministerial Regulation 5 (MR5). However, the laws are quickly warped into mere weapons of authoritarianism as the government uses them to justify censoring journalism and activism.

In this paper, I will argue that these laws on digital censorship are attacks on the ability of CSOs to operate and hold governments accountable for abuses of power. I will first outline the importance of CSOs and their use of the internet to achieve their goals. Second, I examine two of the most prominent censorship laws in Indonesia, the ITE Law and MR5. Next, I discuss recent developments in technology and their impact on digital freedom.

The Monopolization of Information and the Critical Role of Civil Society

Indonesian traditional media is largely incapable of carrying out truthful journalism and informing the public of human rights abuses because of its links to government and government's control of media narratives. The government employs coercive tactics to promote pro-government coverage in the news (Carson and Gibbons 2023). Media outlets who do not fall in line with the government's narratives are threatened with lawsuits for their production of supposed "fake news." Then, legal action dropped in return for more pro-government journalism. For example, in 2018, Hary Tanoesoedibjo, Indonesian media tycoon and owner of 3 corporations that collectively controlled one-third of the nation's viewership, was investigated by the government in a corporate case. Though he previously held a pro-Prabowo stance during the election, Tanoesoedibjo suddenly announced that he would begin supporting Jokowi. Shortly after, the investigations were dropped (Mietzner 2019).

Importantly, the Indonesian media is incredibly monopolized, with a few large conglomerates with vested interests in politics. All of the major media corporations support one political party or another. Surya Paloh, for instance, the owner of Media Group, is the founder of the NasDem political party. Paloh is an ardent supporter of former president Joko Widodo. During past election cycles, Media Group's publications have constantly been pro-Jokowi ("The

Media Landscape in Indonesia: The More Things Change, the More They Stay the Same | FULCRUM" 2022). This gives Indonesia's political elites undue dominance over the content that people consume. Of course, this also means that these huge media conglomerates will never risk publishing content that may tarnish their carefully curated reputation.

Civil society, as Indonesia's only source of credible, government-critical news, plays a crucial role in keeping the Indonesian government accountable for human rights abuses. CSOs uphold a core pillar of democracy and keep voters politically aware. Doing so ensures that all the Indonesian government's misuse of power are known by the public, thus enabling more widespread resistance to problematic actions. For example, the Indonesian government's forced displacement of natives in West Papua was exposed by the CSO TAPOL and subsequently brought to public attention (Human Rights Violations in West Papua Received High Attention in the Recommendations of UN Human Rights Bodies 2024). This accountability mechanism compels officials to be more democratic and improve their governance for fear of public backlash.

To fulfill their broad mission of raising public awareness on state violations of human rights, the internet acts as a campaign tool that allows CSOs to effectively mobilize thousands of members across Indonesia. For example, in 2020, the Government introduced the Omnibus Law on Job Creation. It was widely viewed as an assault on labor rights. In response, 15 CSOs and activist groups organized a national strike with thousands of workers (BBC News 2020). The protests were successful in eliciting concessions from the government when the Supreme Court ruled the bill as unconstitutional and ruled that it had to be amended within 2 years (Indonesian Workers Celebrate Landmark Victory as Constitutional Court Orders Crafting New Labour Legislation | ITUC-Asia Pacific, 2024). In such a large country, the internet and in particular, social media, is perhaps the only way to successfully organize mass protests and strikes.

The Architecture of Digital Repression: ITE Law and MR5

The controversial 2008 Electronic Information and Transaction Law (Undang-Undang Informasi dan Transaksi Elektronik), colloquially known as UU ITE, is at the forefront of Indonesia's digital censorship. The ITE law aims to reduce problematic content on social media by criminalizing misinformation, hate speech, and incitement of violence. This statute is explicitly worded in Article 27 which prohibits the "distribution and transmission" of "insulting or defaming content.". While the law's intentions may have been good, the subjective and overly broad nature of what content constitutes "insults or defamation" is leading to gross miscarriages of justice, especially because of Indonesia's extremely conservative society (Sebastian 2020). In reality, these extremely broad laws punish many ordinary and ultimately harmless actions such as justified criticisms or comments on the government which are key tenets of a healthy democracy.

One prominent example is the 5-year sentence handed to the author Apollinaris Darmawan. Darmawan's arrest was justified by the government because of his "blasphemous behavior "on social media. He posted negative comments on Indonesia's Islamic leaders on the social media platform X (previously Twitter). As a result of his posts, a mob of fanatics forcibly

entered his home and assaulted Darmawan, resulting in a police intervention which eventually ended in his arrest. (Harsono 2024). Far from fulfilling its promise of protecting vulnerable minority populations on the internet, the ITE Law has instead exacerbated the tyranny of majoritarianism. In a clear breach of the human right to freedom of belief, authorities have now been granted legal ground on which to prosecute even mere criticisms against the majority religion of Islam (Jacques 2009). This is especially problematic in Indonesia, where CSOs frequently bring valid criticisms against the majority religion on social issues. CSOs often campaign against frequently misogynistic and homophobic government practices that the government disguises as aspects of Islam doctrine. For example, CSOs such as Komnas Perempuan advocate for the abolishment of abusive and discriminatory dress codes for women while those such as the Setara Institute work to prevent violence against minority faiths in Indonesia. These CSOs are threatened with legal action by the ITE Law because of their anti-government stances.

Furthermore, the government has levied UU ITE directly on CSOs that threaten their interests. This frequently intersects with their desire to silence information about illegal government activity. Notably, in 2022, two activists, Haris Azhar, and Fatia Maulidiyanti, were charged with criminal defamation for speaking out on corruption ("Human Rights Defenders Fatia Maulidiyanti and Haris Azhar, Acquitted of Defamation Charges by the East Jakarta District Court" 2024). The pair were directors of the well-known human rights CSO KontraS. According to officials, their "crime" involved talking about high-ranking military personnel and their links to gold mining enterprises in West Papua on YouTube (Amnesty International 2023). Azhar and Maulidiyanti argued that the Indonesian military was protecting their private interests in the war-torn region. Though the charges against Azhar and Maulidiyanti were eventually dropped, their case has served as a dangerous precedent on the extent to which Indonesian authorities are willing to silence civil society.

Digital censorship has evolved beyond the ITE law, extending its reach into the private sector. In 2020, a new bill called Ministerial Regulation 5 (MR5) was introduced. Its purpose was to require social media giants such as Google, Twitter, and Facebook to comply with government orders to expunge content from their platforms on government demand. MR5 included clauses that mandate these platforms hand over personal data of its users to the government upon request (Strangio 2022). The glaring issue with MR5 is that it has made identifying, tracking, and punishing people on the internet extremely easy for the government. MR5 has deanonymizing the digital realm for the Indonesian government. Anonymity is a cornerstone of the internet's role as a forum for free speech because it allows people, in particular minorities, to speak their mind without the fear of repercussion (York 2022).

Since the implementation of MR5, whistleblowers are far more hesitant to expose important information on government misdeeds because their online activity can be traced back to them (REF). For example, in 2020, the police arrested seven administrators of a WhatsApp group, a Facebook page, and an Instagram account used to mobilize and organize protests against the Omnibus Law on Job Creation mentioned above (Rahma 2020). They were part of the labor

CSO "Panjang Umur Perlawanan." Their arrests signal a brutal suppression of the democratic right to protest.

Furthermore, companies have to fulfill these orders no less than 24 hours after they have been made by the government, lest they risk their service being completely blocked in Indonesia (Strangio 2022). Nearly 9,000 tech companies have already agreed to MR5, severely jeopardizing the right to freedom of expression in Indonesia. This complicity can be attributed to corporations' profit-driven decision-making. With a tech industry valued in the tens of billions of dollars and hundreds of millions of potential customers, Indonesia is a lucrative market for tech platforms (Negara and Meilasari-Sugiana 2022). Since they are unwilling to lose the opportunity to capitalize on the market, they then comply with MR5.

Systematic Undermining of Civil Society

Pro-censorship proponents claim that the Indonesian government simply intends to eliminate hate speech and reduce harassment online, making the internet safer. Though there do exist provisions in the ITE Law and MR5 that are meant to shield activists and minorities from online hate speech, these provisions have largely failed to be implemented. In 2024, The Ministry of Communication and Information Technology revealed that across social media platforms, hate speech against ethnic and religious minorities has increased by 30% when compared to 2023, despite the new laws whose supposed aim is to decrease this (Navigating the Digital Minefield: Insights from PowerTalk on Hate Speech, 2024). This is of no surprise because Indonesia's political sphere has long been dominated by Islamist majoritarianism. Minority rights are severely unprotected because there still exists many draconian Blasphemy Laws in Indonesia. When minorities face backlash, such as death threats and insults for criticizing the majority religion, the pro-government perpetrators are rarely punished. Mostly, the primary victims of censorship are activist groups as evidenced by the fact that in 2018, journalists ranked as one of the most targeted professionals by the ITE law (Carson and Gibbons 2023b). Therefore, in the vast majority of instances, censorship laws are used to punish government dissenters instead of digital trolls and haters.

Despite repeated condemnations from the international community, there is no end in sight to Indonesia's draconian ITE law. Instead, the government seems to have only become bolder and more aggressive in its restrictions. This can be observed in the sheer quantity of ITE Law-related trials, which has more than tripled from 74 during Susilo Bambang Yudhoyono's second term (2009-2014) to 233 during Joko Widodo's first term (2014-2019). Even more jarringly, the officials are continuously lowering the bar for what qualifies as a violation of the law, and they are becoming increasingly punitive. For example, during the Widodo Administration, of the 233 ITE cases, 82 were tried for insulting the president (Hamid 2019). Yet, there is no consistency on what "insulting the president" means. One of these cases even involved an 18-year-old student who in 2018, was arrested for sharing unflattering memes about the president on Facebook (BBC News, 2018). In what is now becoming a vicious negative feedback loop, the government's digital repression has made it ever more politically unpopular

amongst its own people. This has given rise to resistance from CSOs, which with dystopian regularity, triggers heavy-handed crackdowns from authorities.

Now that the government has demonstrated unprecedented ruthlessness in pursuing its commitment to full digital repression through ITE and MR5, CSOs are more hesitant than ever to speak out. A recent survey from INDIKATOR (2020) shows that almost 70% of Indonesians reported that they were afraid to openly express their opinions. When people live in fear of engaging in activism, this greatly reduces the power and capacity of CSOs to pursue impactful societal change. Historically, CSOs have relied on passionate and eager volunteers to carry out important roles within their organizations (Young, 2023). Now that these very people are seeing the punishments levied against anti-government activists, it prevents them from speaking freely lest they become the next target (Post, 2020). CSOs themselves have also echoed that their survival is in danger because of the lack of people engaging with them. A survey of 62 CSO leaders from across Indonesia revealed that all CSOs face threats from the state and state-sponsored actors (2023) because; the ITE law and MR5 has made speaking out on human rights incredibly difficult for CSOs (Robet et al., 2023). These laws are coupled with the government increasingly capitalizing on technological advances to strengthen their grip on Indonesian cyberspace by developing more sophisticated methods of digital censorship. In 2018, Indonesia's Ministry of Communication and Information Technology upgraded their capacities by purchasing a new web crawler that was to be operated by a specialist team called Cyber Drone 6 (Gov't Launches "Web Crawler" to Seek out Negative Internet Content, n.d.). This has essentially automated the process of detecting undesirable content on social media, making the government's censorship capabilities even more powerful. If Indonesia's ongoing digital authoritarianism persists, the future of civil society is bleak.

Conclusion

In conclusion, Indonesia's recent digital censorship laws directly threaten the survival and efficacy of CSOs. First, it prevents people from knowing about the government's human rights abuses which directly reduces democratic accountability. So long as their crimes are not brought to light, they will always continue. Second, censorship creates a chilling effect which makes people afraid to engage in civil society. This drives CSO volunteers away and makes it harder for existing activists to safely continue doing their work. Finally, censorship is being disproportionately weaponized against religious minority groups, which violates the very purpose of its existence. These harms are preventing Indonesia's democratization by blocking the right of ordinary citizens to exercise their political power based on truthful information. There is great urgency to reform Indonesia's digital censorship laws.

Works Cited

- Amnesty International. 2023. "Indonesia: Further Information: Human Rights Defenders Charged With Defamation: Haris Azhar and Fatia Maulidiyanti - Amnesty International." January 18, 2023. <https://www.amnesty.org/en/documents/asa21/5461/2022/en/>.
- BBC News. 2018. "Indonesia Teenager Jailed for President Widodo Insults." January 16, 2018. <https://www.bbc.com/news/world-asia-42710115>.
- BBC News. 2020. "Indonesia: Thousands Protest Against 'Omnibus Law' on Jobs." October 8, 2020. <https://www.bbc.com/news/world-asia-54460090>.
- Bertrand, Jacques. "Political Islam and Democracy in the Majority Muslim Country of Indonesia." In *Islam and Politics in Southeast Asia*, pp. 61-80. Routledge, 2009.
- Bischoff, Paul, and Paul Bischoff. 2023. "Internet Censorship 2024: A Map of Internet Censorship and Restrictions." Comparitech. October 16, 2023. <https://www.comparitech.com/blog/vpn-privacy/internet-censorship-map/>.
- Carson, Andrea, and Andrew Gibbons. 2023. "The Big Chill? How Journalists and Sources Perceive and Respond to Fake News Laws in Indonesia and Singapore." *Journalism Studies* 24 (14): 1819–38. <https://doi.org/10.1080/1461670x.2023.2192299>.
- Davis, Ben. "Financial Sustainability and Funding Diversification: The Challenge for Indonesian NGOs." *International Journal of Voluntary and Nonprofit Organisations* 5 (2013): 1-26.
- Freedom House. 2024. *The Struggle for Trust Online*. Freedom House. <https://freedomhouse.org/report/freedom-net/2024/struggle-trust-online>.
- "Gov't Launches 'Web Crawler' to Seek out Negative Internet Content." Jakarta Globe. <https://jakartaglobe.id/news/govt-launches-web-crawler-to-seek-out-negative-internet-content>.
- Hamid, Usman. 2019. "Indonesia's Information Law Has Threatened Free Speech for More Than a Decade. This Must Stop." *The Conversation*. <https://theconversation.com/indonesias-information-law-has-threatened-free-speech-for-more-than-a-decade-this-must-stop-127446>.
- Harsono, Andreas. 2024. "Behind Bars for 'Blasphemy' in Indonesia." Human Rights Watch, January 10, 2024. <https://www.hrw.org/news/2023/12/23/behind-bars-blasphemy-indonesia>.
- Henshall, Will. 2023. "Global Internet Freedom Declines, Aided by AI." *TIME*, October 4, 2023. <https://time.com/6319723/global-internet-freedom-decline-2023/>.
- "Human Rights Defenders Fatia Maulidiyanti and Haris Azhar, Acquitted of Defamation Charges by the East Jakarta District Court." 2024. Front Line Defenders, July 25, 2024. <https://www.frontlinedefenders.org/en/case/human-rights-defenders-fatia-maulidiyanti-and-haris-azhar-acquitted-defamation-charges-east>.
- "Indonesia: Number of Internet Users 2028 | Statista." 2023. Statista, December 12, 2023. <https://www.statista.com/statistics/254456/number-of-internet-users-in-indonesia/>.
- "Indonesian Workers Celebrate Landmark Victory as Constitutional Court Orders Crafting New Labour Legislation." 2024. ITUC-Asia Pacific. <https://www.ituc-ap.org/news-and-updates/indonesian-workers-celebrate-landmark-victory-as-constitutional-court-orders-crafting-new-labour-legislation>.

- Mietzner, Marcus. 2019. "Authoritarian Innovations in Indonesia: Electoral Narrowing, Identity Politics and Executive Illiberalism." *Democratization* 27 (6): 1021–36.
<https://doi.org/10.1080/13510347.2019.1704266>.
- "Navigating the Digital Minefield: Insights from PowerTalk on Hate Speech." 2024. Monash University, Indonesia, August 15, 2024.
<https://www.monash.edu/indonesia/news/navigating-the-digital-minefield-insights-from-powertalk-on-hate-speech-monitoring-and-online-polarisation>.
- Negara, Siwage Dharma, and Astrid Meilasari-Sugiana. "The State of Indonesia's Digital Economy in 2022." *FULCRUM*, November 23, 2022.
<https://fulcrum.sg/the-state-of-indonesias-digital-economy-in-2022/>.
- Post, J. 2020. "Growing Fear of Speaking Out: Survey Finds Indonesia's Civil Liberties under Threat." *The Jakarta Post*, October 26, 2020.
<https://www.thejakartapost.com/news/2020/10/26/growing-fear-of-speaking-out-survey-finds-indonesias-civil-liberties-under-threat.html>.
- Rahma, Andita. "Police Arrest 7 Social Media Administrators Linked to Omnibus Law Protest." *Tempo*, October 20, 2020.
<https://en.tempo.co/read/1397594/police-arrest-7-social-media-administrators-linked-to-omnibus-law-protest>.
- Robet, R., Fitri, M. R., and Kabelen, M. C. S. 2023. "The State and Human Rights under Joko Widodo's Indonesia." *Cogent Social Sciences* 9 (2).
<https://doi.org/10.1080/23311886.2023.2286041>.
- Sebastian, Leonard C., Syafiq Hasyim, and Alexander R. Arifianto, eds. *Rising Islamic Conservatism in Indonesia: Islamic Groups and Identity Politics*. Routledge, 2020.
- Strangio, Sebastian. 2022. "Deadline Looms for Indonesia's Harsh New Internet Content Restrictions." *The Diplomat*, July 19, 2022.
<https://thediplomat.com/2022/07/deadline-looms-for-indonesias-harsh-new-internet-content-restrictions/>.
- "The Media Landscape in Indonesia: The More Things Change, the More They Stay the Same." 2022. *FULCRUM*, November 29, 2022.
<https://fulcrum.sg/the-media-landscape-in-indonesia-the-more-things-change-the-more-they-stay-the-same/>.
- Top 10 Media & Internet Manufacturing Companies in Indonesia by Revenue for November 2024. *ZoomInfo*, 2024.
<https://www.zoominfo.com/top-lists/top-10-companies-from-media-industry-in-ID-by-revenue>.
- WJP Rule of Law Index. 2024. *Global Insights*.
<https://worldjusticeproject.org/rule-of-law-index/global/2024/Indonesia/Absence%20of%20Corruption/>.
- York, Jillian C. 2022. "The Right to Anonymity Is Vital to Free Expression: Now and Always." *Electronic Frontier Foundation*, January 25, 2022.
<https://www.eff.org/deeplinks/2020/03/right-anonymity-vital-free-expression-now-and-always>.
- Young, W. 2023. "Volunteering as a Driving Force for NGOs." *NGOReach*, August 2, 2023.
<https://www.ong-ngo.org/volunteering-as-a-driving-force-for-ngos/>.

From Concealment to Confrontation: France's Shifting Memory of Nazi Occupation and the Algerian War By Jean-Louis Andreani

Abstract

This paper explores France's evolving collective memory concerning the Nazi Occupation and the Algerian War, focusing on the political and institutional responses that have shaped national identity. It argues that while France has made strides in acknowledging its role in past atrocities, these efforts often fall short of genuine reconciliation due to selective memory and the absence of formal apologies. The essay critically examines the approaches of key French leaders—Charles de Gaulle's and François Mitterrand's emphasis on unity at the expense of confronting complicity, Jacques Chirac's pivotal acknowledgment of Holocaust collaboration, and the limited advances under Nicolas Sarkozy and Emmanuel Macron in addressing the legacy of the Algerian War. Through an analysis of these historical reckonings, the paper contends that France's nation-building strategies have perpetuated incomplete narratives, thereby impeding true reconciliation. The essay concludes by advocating for comprehensive apologies and the integration of marginalized histories as essential steps toward a cohesive and honest national memory.



11

Figure 1: The Deportation of Jews From France

Introduction

In yet another iteration of the “banality of evil,”¹² this haunting image captures the grim collaboration between German and French policemen as they deport Jewish families during the Nazi occupation of France. This stark reminder of one of France's darkest chapters raises

¹¹ Figure 1

¹² Bernard J. Bergen, *The Banality of Evil* (Lanham, MD: Rowman & Littlefield, 2000).

profound questions about memory, responsibility, and the narratives we choose to remember—or forget. Of particular interest here is how the collective processes of remembering and forgetting are orchestrated and directed by political institutions and state-based actors. In contemporary France, the legacy of the Nazi occupation during World War II (1940-1944) and the brutalities of the Algerian War (1954-1962) remain deeply embedded in the nation's collective consciousness, yet also disavowed - collective traumas were simultaneously registered and repudiated. Hence, for many years, the French government has attempted to obscure its complicity in these atrocities, promoting narratives that downplayed or refused *tout court* the darker aspects of what transpired. However, the rewriting of history as the repression of traumatic memory rarely fully succeeds - following Freud's topology, "the return of the repressed" remains forever smashed through personal and collective defense mechanisms. Stated otherwise, these two tragedies loomed silently in national memory, waiting to be "heard," "spoken," and truly "exorcized." It took France many years, even decades, to start facing the uncomfortable truths about its past. Yet, while it began to address these issues, it avoided fully confronting them and often used biased political language to discuss these national traumas without truly engaging with them. This essay asserts that French government initiatives have shaped modern France's collective memory and identity to construct a hegemonic narrative that may lend itself to nation-building, potentially at the price of the truth.

The Nazi Occupation

During the Nazi occupation of France, the Vichy government collaborated with the Third Reich and persecuted Jews. French police assisted in the roundup and deportation of around 75,000 Jews¹³, most of whom were sent to concentration camps like Auschwitz, where they were murdered and gassed. The Vel' d'Hiv Roundup in 1942 was a notorious example, where over 13,000 Jews were arrested in Paris alone¹⁴. Ever since the war, France has been grappling with its role in these atrocities. At first, it only punished those it deemed responsible¹⁵. Only much later did it come to accept its mistakes as a united nation.¹⁶

The devastation of war left profound wounds and lasting trauma on the French. Around 400,000 buildings were destroyed, and five times that number was damaged. Industrial and agricultural production was running at just 40% of what it had been pre-war. The French population was sick and starving: rationing would continue until 1949, and two-thirds of children

¹³ "Le Bilan de la Shoah en France [Le régime de Vichy]," *bseditions.fr*, www.bseditions.fr/article.php?pArticleId=158&pChapitreId=23982.

¹⁴ "Pourquoi la rafle n'a pas atteint son objectif." *AIDH.org*, p. 52, <http://aidh.org/pdf/rafle-obstacle>.

¹⁵ Jean-Paul Cointet, "Épuration légale: 400 000 dossiers, moins de 800 morts," *Historia* (in French).

¹⁶ Given the vast and complex nature of France's grappling with its wartime role, this paragraph offers only a brief summary of pivotal moments and does not capture the full scope of the nation's reconciliation efforts. For more resources, refer to Henry Rousso's *The Vichy Syndrome: History and Memory in France since 1944* (Harvard University Press, 1991), Robert Paxton's *Vichy France: Old Guard and New Order, 1940-1944* (Columbia University Press, 1972) and Sarah Gensburger's *Witnessing the Robbing of the Jews: A Photographic Album, Paris, 1940-1944* (Indiana University Press, 2015).

were suffering from rickets. One child in 10 did not survive childbirth¹⁷¹⁸. While citizens who participated in the Resistance were celebrated and recognized, those who embraced Vichy's regime were punished: people suspected of collaborating were executed by the government, and women who had had relations with German soldiers had their heads shaved. In the middle of this chaos lay two conflicting forces: on the one hand, Charles de Gaulle's desire to craft a narrative of a triumphant, unified France emerging from the shadows of occupation - a new Republic that would defeat the specters of recent history. On the other, the everyday experiences of French citizens, who struggled to reconcile their memories of suffering, loss, and complicity. De Gaulle's focus on liberation overshadowed the darker aspects of the war, as he sought to close the door on a painful past. After France was liberated, De Gaulle allowed his countrymen to witness the German surrender in an attempt to put France "in a position to [...] go on" ("this [...] allowed the French to evade what had actually happened during the war—the abject armistice with Hitler, the delivery of the deportees to their persecutors, the entire black hole of Vichy from which so little light escaped for so long")¹⁹²⁰. This attitude of propping up the good and ignoring the bad effectively misled the population into subconsciously forgetting the atrocities of its country, and it persisted for some time. It was replicated by De Gaulle again when he ran for president of the Fifth Republic in 1958, and by Nicolas Sarkozy much later in 2007. But is this truly enough of a reckoning, or does it shift the blame onto those actors without addressing the deeper source of the problems?

While some argue that France's postwar trials were significant steps toward accountability, they fell short of a full reckoning. These trials, while seen as efforts to bring justice, ultimately served to scapegoat a few individuals, allowing France to distance itself from its trauma. By focusing on punishing a select group, the nation avoided confronting the deeper, systemic issues and collective responsibility that enabled the Vichy regime to thrive (and also projected its trauma onto those individuals - purging the agony - catharsis, etc.).

The French disavowal of Vichy is systemic and thus we must also examine how, during the twenty-five years following the war, the French government dismissed the horrors of the Nazi deportation of Jews, choosing instead to focus on narratives that absolved or minimized their own complicity, such as the glorification of the Resistance²¹²². After the Second World War came to a close, the French felt extreme guilt and shame for having colluded with the Nazis and

¹⁷ Mark Davis, "How World War II Shaped Modern France," *Euronews*, May 5, 2015, <https://www.euronews.com/2015/05/05/how-world-war-ii-shaped-modern-france>.

¹⁸ George W. Kyte, "War Damage and Problems of Reconstruction in France, 1940-1945," *Pacific Historical Review*, vol. 15, no. 4, 1946, pp. 417–26, <https://doi.org/10.2307/3635778>.

¹⁹ Adam Gopnik, "How Charles de Gaulle Rescued France," *The New Yorker*, August 13, 2018, <https://www.newyorker.com/magazine/2018/08/20/how-charles-de-gaulle-rescued-france>.

²⁰ Julian Jackson, *De Gaulle* (Cambridge, MA: Harvard University Press, 2018).

²¹ Marlise Simons, "Chirac Affirms France's Guilt in Fate of Jews," *The New York Times*, July 17, 1995, <https://www.nytimes.com/1995/07/17/world/chirac-affirms-france-s-guilt-in-fate-of-jews.html>.

²² Tsilla Hersheo, "At Issue: The Jewish Resistance in France during World War II: The Gap between History and Memory," *Jewish Political Studies Review*, vol. 19, no. 1/2, 2007, pp. 49–57, <http://www.jstor.org/stable/25834722>.

were reluctant to confront this truth²³. This denial came to be known as Vichy Syndrome²⁴; coined by historian Henry Rousso, it depicts “an agonized postwar France somehow attempting to reconcile itself to its history”²⁵. Indeed, when De Gaulle became the first president of the Fifth Republic in 1958, his narrative of “a nation unified against the Nazi oppressor” was characterized by Professor Nathan Bracher as containing “distortions and lacunae”, and his account of French grandeur was “highly selective”²⁶. Nevertheless, he was a respected figure—and rightfully so—he was, after all, considered a French war hero—and thus received support from a sizable majority. The resistance narrative prevailed, yet, there was still a concerning silence around the Vichy collaboration with the Nazis, especially when compared to those surrounding the acts of the Resistance.

The emphasis on the heroic hagiography of the Resistance served as a convenient way to sidestep the uncomfortable truths of the Vichy collaboration. It was as if the government had decided that if it were going to revisit this period, it should be done through the lens of the Resistance’s valor, making it the cornerstone of the collective consciousness.

Similarly, when former president François Mitterand was elected in 1981, he separated the Vichy Regime from the French Republic. He refused to apologize, claiming that “the Republic had nothing to do with [the Vichy Regime]” and that France was not responsible for any crimes committed²⁷. Moreover, he also refused to “make a formal declaration of recognition”²⁸. In an interview with Jean-Pierre Elkabbach²⁹, Mitterand doubled down on his position, asserting that France would not present any apologies (“La France n’a pas d’excuses à donner”). He even attacked those who wanted an apology or recognition, saying that it was an excessive demand from people who did not profoundly feel the honor of being French. This hyperbolic response can be understood as a defensive maneuver to protect the sanctity of French Republican history. By framing calls for acknowledgment as unpatriotic, Mitterrand deflected scrutiny away from deeper questions about France’s national identity and the Republic’s moral failings during the Vichy era. His stance is also likely influenced by his own controversial connections to Vichy; a duality that raised persistent questions about his loyalties and motivations. This personal history may have made him particularly resistant to framing Vichy’s actions as an integral part of French history, as it risked implicating him and undermining his legacy. By doing this, Mitterand implies that those wanting an apology do not resonate with

²³ See more on German Collective Guilt after WWII): Jaspers, Karl, and E. B. Ashton. *The Question of German Guilt*. New York: Fordham University Press, 1965. <https://doi.org/10.2307/j.ctt13wzz9w>.

²⁴ Full Definition: Scullion, Rosemarie. 1999. *Unforgettable: History, Memory, and the Vichy Syndrome*. *Studies in 20th & 21st Century Literature* 23 (1) p. 1. <https://doi.org/10.4148/2334-4415.1452>.

²⁵ Bertram M. Gordon, “The ‘Vichy Syndrome’ Problem in History,” *French Historical Studies*, vol. 19, no. 2, 1995, pp. 495–518, <https://doi.org/10.2307/286785>.

²⁶ Nathan Bracher, “Bruckner and the Politics of Memory: Repentance and Resistance in Contemporary France,” *South Central Review*, vol. 24, no. 2, 2007, pp. 54–70, <https://www.jstor.org/stable/40039981>.

²⁷ Marlise Simons, “Chirac Affirms France’s Guilt in Fate of Jews,” *The New York Times*, July 17, 1995, <https://www.nytimes.com/1995/07/17/world/chirac-affirms-france-s-guilt-in-fate-of-jews.html>.

²⁸ Peter Carrier, *Holocaust Monuments and National Memory Cultures in France and Germany Since 1989* (New York: Berghahn Books, 2005).

²⁹ Jean-Pierre Elkabbach, *François Mitterrand, Conversation avec un Président*, documentary, 1994.

France, and by the same token the actions of the Vichy Regime do not align with France. However, no matter what he claimed, Vichy would always be ingrained in the history of the Republic, and denying this fact demonstrates the extent to which politicians feel they can choose what, or what not, to include in the story, even though it had already happened. This is a powerful quote from Mitterrand, as it truly depicts the government's grasp on the "story," as he outright denies the existence of real events.

Mitterrand's actions during the Vichy regime are a very controversial subject, with many believing that he was an ally to the government despite reports of his resistance behavior. This suspicion lends credence to the idea that his active rejection of Vichy later in his career was not just a reflection of national reconciliation efforts but also a strategic attempt to distance himself from his own associations with the regime. According to Franz-Olivier Giesbert, Mitterrand "came to Vichy to work"³⁰ and discovered letters revealing that he qualified Pétain as "magnificent", comparing his face to "that of a marble statue"³¹. Furthermore, to the outrage of the Jewish community, he would often stop by Pétain's grave and adorn it with flowers: on the 22nd of September, 1984; the 15th of June, 1986; and every 11th of November from 1987 to 1992³². This tradition only stopped after heavy protests from the Jews³³, and Mitterrand explained that he was simply following what previous presidents had done before³⁴. Finally, the straw that broke the camel's back involves Mitterrand's friend René Bousquet. The latter was a general secretary of the Vichy police and their friendship aroused many questions: Lionel Jospin, former prime Minister of France from 1997 to 2002, was perplexed by "[Mitterrand's] continuing relationship into the 80s with the likes of Bousquet who organized the mass arrests of Jews"³⁵; Pierre Moscovici, another French politician, was "shocked" by Mitterrand's friendship with "someone [...] instrumental in state antisemitism and the final solution", claiming that René Bousquet was "absolute evil"³⁶. When put all together, there is a solid case regarding Mitterrand's connections with the Vichy regime, but we must also remember that these are rumors and take them with a grain of salt. Regardless, Mitterrand's trustworthiness is suspect, adding a hovering question mark over his words.³⁷

After a generation of presidents³⁸ failed to acknowledge France's collusion with the Nazis, it was finally former president Jacques Chirac who broke this tradition when he was

³⁰ *François Mitterrand ou la tentation de l'histoire*, Éditions du Seuil, 1977, ISBN 2-02-004591-5, chap. 5, pp. 43–50.

³¹ Pierre Péan, *Une Jeunesse Française* (Paris: Fayard, 2014).

³² Henry Rousso, *Le syndrome de Vichy, de 1944 à nos jours* (Paris: Éditions du Seuil, coll. « Points histoire », 1990), p. 389.

³³ "Vie-publique.fr," archived April 12, 2008, at the Wayback Machine, extracts from the interview François Mitterrand gave to Radio J on November 13, 1992.

³⁴ *La Décennie Mitterrand* (éd. du Seuil, coll. « Points », 2001), pp. 646–647.

³⁵ "FRANÇOIS MITTERRAND Répondra Aux Questions de Jean-Pierre Elkabbach," *L'Humanité*, September 11, 1994, <https://www.humanite.fr/-/francois-mitterrand-repondra-aux-questions-de-jean-pierre-elkabbach>.

³⁶ "FRANÇOIS MITTERRAND Répondra Aux Questions," *L'Humanité*, September 11, 1994.

³⁷ Note: rumors

³⁸ For the purposes of this paper, we are focusing on the discourse of De Gaulle and Mitterrand as they were the most engaged with the subject See Hollande, Pompidou, and Valéry D'Estaing for further information.

elected in 1995. Chirac publicly recognized the role the nation played in the oppression of Jews. He claimed that those dark hours were an injury to France's past and traditions ("[ces heures noires] sont une injure à notre passé et à nos traditions") and that they will forever stain the country's history ("ces heures noires souillent à jamais notre histoire")³⁹. He then went on to give specifics, detailing how 450 French policemen and gendarmes complied with the Nazis to arrest nearly ten thousand Jewish men, women, and children⁴⁰. Chirac condemns this catastrophe, labeling it as irreparable ("[La France accomplissait] l'irréparable") and asserts that France has failed to protect its people and instead delivered them to their deaths ("Manquant à sa parole, [La France] livrait ses protégés à leurs bourreaux"). The country did not live up to its status of "terre d'accueil et d'asile" (land of welcome and asylum), instead failing to defend those it vowed to protect.

Nevertheless, Chirac's acknowledgment marked a significant turning point in the national consciousness. His public admission forced the nation to confront its complicity and the moral failures that accompanied it⁴¹. At the time, newspapers ranging from *Le Monde* to *The New York Times* praised Chirac, noting that he "went well beyond the positions of his predecessors."⁴² Renowned Nazi hunter Serge Klarsfeld hailed Chirac for his "courage,"⁴³ while Henri Hajdenberg, president of the Representative Council of Jewish Institutions of France, described the moment as a turning point: "It is, finally, looking the truth in the face, lifting the veil."⁴⁴ Joseph Sitruk, the chief rabbi of Paris, expressed that he was "fully satisfied"⁴⁵ with the statement, and Jean Kahn, president of the European Jewish Congress, remarked that Jews, as well as all those who fought the Nazis, "must have been delighted to hear these words."⁴⁶ But, going back to Sitruk's words, can one truly be "fully satisfied" with mere acknowledgment? Of course, recognizing the darkness is the first step but the larger historical and psychological question remains: what does it mean to be fully satisfied in the face of such profound historical wrongs? Still, by bringing these dark chapters to light, Chirac initiated a broader conversation about accountability, memory, and justice, signaling a shift toward greater transparency and responsibility. His speech not only challenged the narrative of French innocence but also set the

³⁹ "Allocution de M. Jacques Chirac Président de la République prononcée lors des cérémonies commémorant la grande rafle des 16 et 17 juillet 1942 (Paris)" (PDF), *Jacqueschirac-asso* (in French), July 16, 1995, www.jacqueschirac-asso.

⁴⁰ "Allocution de M. Jacques Chirac," www.jacqueschirac-asso.

⁴¹ It is important to note that academics were discussing these things previous to Chirac's comments: see Robert Paxton (curiously enough, he was the first historian to write about this topic even though he is American)

⁴² Marlise Simons, "Chirac Affirms France's Guilt in Fate of Jews," *The New York Times*, July 17, 1995, <https://www.nytimes.com/1995/07/17/world/chirac-affirms-france-s-guilt-in-fate-of-jews.html>.

⁴³ "Chirac Admits France's Complicity with Nazis: Europe: For the First Time, a French President Acknowledges the Nation's Role in the Deportation of Jews during World War II," *Los Angeles Times*, July 17, 1995, <https://www.latimes.com/archives/la-xpm-1995-07-17-mn-24867-story.html>.

⁴⁴ "Chirac Admits France's Complicity with Nazis," *Los Angeles Times*, July 17, 1995.

⁴⁵ Marlise Simons, "Chirac Affirms France's Guilt in Fate of Jews," *The New York Times*, July 17, 1995, <https://www.nytimes.com/1995/07/17/world/chirac-affirms-france-s-guilt-in-fate-of-jews.html>.

⁴⁶ Marlise Simons, "Chirac Affirms France's Guilt," *The New York Times*, July 17, 1995.

stage for ongoing efforts to address the wrongs of the past, making it a crucial step in the country's journey toward historical reconciliation.

While Chirac's acknowledgment of France's complicity marked a turning point toward greater transparency, subsequent leaders, like President Nicolas Sarkozy, continued to emphasize narratives that celebrated the Resistance, often downplaying the darker aspects of French collaboration with the Nazis. At the inaugural commemoration in the Bois de Boulogne, the newly elected president, following in De Gaulle's footsteps, wanted to acknowledge the Resistance's military and political triumph. To do so, he decided to focus national attention on 35 resistants who valiantly died fighting the Nazis, appointing Max Gallo to deliver an eloquent speech ("These 35 heroes are our present. They show us the courage and spirit of sacrifice of young people when they are uplifted by the power of idealism and love for the nation")⁴⁷. However—and this is where Sarkozy's approach differs from De Gaulle's—the resistants were not only promoted as a symbol of heroism but also as a symbol of uniting through diversity, putting differences aside to strive for a greater good. All thirty-five resistants had "differing social origins and political sensitivities" but shared "a common goal: to make France free and sovereign so that its people will control their own destiny"⁴⁸. This powerful assertion not only presented the Resistance as heroic but also established it as a model of unity and accepting diversity for the nation to follow⁴⁹. Unlike De Gaulle, Sarkozy was willing to recognize the country's faults. Gallo was quick to highlight the collaboration between the French and Nazis: "There were indeed French citizens willing to betray, torture, and massacre"⁵⁰. But, he refuses to blame the entire French nation for these actions. Instead, he believes that the true "soul of the nation" is embodied by the young resistants who fought to see Paris—described by de Gaulle as "violated and broken"—eventually "liberated by itself" and "by its people"⁵¹.

While Sarkozy acknowledged the collaboration between the French and the Nazis, he refused to credit this collaboration to the entire nation, instead selecting a few individuals who, according to him, did not represent the state. This is disingenuous, as it not only fails to exonerate the French but also raises the question of how these "few individuals" gained the necessary support to enable such collaboration. As such, this recognition was insufficient and should have been accompanied by deeper, more comprehensive repentance. The mention of the collusion with Germany was primarily a clever trick to reinforce values of patriotism and heroism throughout the nation. It almost undermines this goal by depicting those who participated in the collaboration as not truly representative of France, instead emphasizing that the real embodiment of the nation's spirit were those who fought to liberate the country. Yet it is undeniable that those who participated in the collaboration were French. They did not represent

⁴⁷ Nathan Bracher, "Bruckner and the Politics of Memory: Repentance and Resistance in Contemporary France," *South Central Review*, vol. 24, no. 2, 2007, pp. 54–70, <https://www.jstor.org/stable/40039981>.

⁴⁸ Nathan Bracher, p. 66.

⁴⁹ It is ironic to instrumentalize the Resistance in a moment where French debates about French multiculturalism are live and complex

⁵⁰ Nathan Bracher, p.67

⁵¹ Nathan Bracher, p. 67

French values, such as liberty, equality, and fraternity, but the reality remains: they were French. This highlights a deeper issue—both then and now, there is no clear consensus on what France truly represents. The narrative of France is not singular; it is complex, branched, and conflicting.

France's collective memory of the Vichy collaboration has evolved into an incomplete reconciliation. While leaders like Jacques Chirac took significant steps by acknowledging the nation's complicity, others have chosen to downplay or reinterpret this history for political ends. These selective narratives demonstrate the ongoing struggle within France's memory politics, oscillating between unity and accountability. Still, Chirac's acknowledgment, though imperfect, marked a pivotal moment, highlighting the importance of confronting difficult truths. This evolving memory reflects the unresolved tensions between the ideals of the French Republic and the realities of its past.

The first step of authentic reconciliation requires acknowledging the actions of those complicit and the broader implications for the nation. France's reckoning with its collaboration during World War Two demonstrates that acknowledgment, however delayed or imperfect, is a vital foundation for addressing historical injustices. Once this initial step is taken, society can begin enacting meaningful reforms, such as educational initiatives, policy changes, and public discourse that address historical injustices. The reckoning with collaboration revealed the dangers of crafting selective narratives to avoid uncomfortable truths. Future efforts must approach these truths with transparency and accountability. The question remains: what should this "confrontation" look like?

The Algerian War

The Nazi collaboration was not the only major event in French history that the country was ashamed of and avoided for many years. The Algerian War holds an equally shameful place in the nation's memory due to the numerous war crimes committed by the French, as well as the controversy that emerged from its aftermath and the moral questions it raised about colonialism, violence, and national identity. The war's death toll is uncertain, ranging from 400,000 to 1.5 million Algerians^{52,53}. The French committed various war crimes, including rape, massacres of innocent civilians, and torture. Dr. Kevin Shillington reports that "more than 8,000 villages had been destroyed in the fighting", roughly "three million people were displaced" and "one million Algerians and some 10,000 colons lost their lives"⁵⁴. There were around "hundreds of thousands of instances of torture" committed by the French⁵⁵. Additionally, internment camps were also used during the war⁵⁶ as well as after the war to detain Harkis (Algerians who fought with the

⁵² "Ombres et lumières de la révolution algérienne," *Le Monde diplomatique*, November 1, 1982.

⁵³ FRANCE 24, "France Remembers the Algerian War, 50 Years On," *France 24*, March 16, 2012, <https://www.france24.com/en/20120316-commemorations-mark-end-algerian-war-independence-france-evian-accords>.

⁵⁴ Kevin Shillington, *Encyclopedia of African History: 3-Volume Set* (New York: Routledge, 2013).

⁵⁵ Pierre Vidal-Naquet, *Les crimes de l'armée française: Algérie, 1954-1962* (La Découverte, November 20, 2014), p. 118, ISBN 978-2-7071-8309-5.

⁵⁶ Marc Bernardot, *Camps d'étrangers* (in French) (Paris: Terra, 2008).

French)⁵⁷.

The trauma of the Algerian War extends beyond the battlefield, seeping into the very fabric of French society. For soldiers who served in Algeria, the experience often left deep psychological scars, as many were involved in or witnessed acts of brutality. Bachir Hadjadj, a fighter in the National Liberation Army (ALN), the military arm of the National Liberation Front (FLN) that fought for Algerian independence, recounted how he felt extreme disgust: “I saw torture. I heard how the soldiers treated villagers. Surrounding them, raiding their villages, exposing women’s genitals to see if they were shaved [...] I felt resentment.”⁵⁸ Similarly, Serge Carel, a harki⁵⁹, commented on an act of revenge against innocent villagers, which marked a moment of disillusionment with the actions of his side: “That shocked me.” He also conveyed the deep trauma and horror he endured during the war, expressing his regret with the words: “This period of time fills me with horror.”⁶⁰ Finally, Jacques Massu, a former French general who served during this time, expressed his remorse over the use of torture⁶¹. Revisiting these experiences was shameful and challenging after the war, and the massive influx of Algerian immigrants into France only deepened the unresolved wounds of the past.

Moreover, the war aggravated tensions within France, as the influx of Algerian immigrants heightened social and cultural divisions. The number of Algerians in France skyrocketed, going from 211,000 in 1954 to 350,000 in 1962⁶². The Harkis were subject to memory politics “on both sides of the Mediterranean”⁶³: in Algeria, they were vilified as traitors to the independence movement, while in France, they were marginalized and largely excluded from the dominant national narrative, further entrenching their sense of alienation. As Benjamin Stora asserts, while France engaged in a forty-year policy of deliberately forgetting about the conflict, Algeria mythologized the war as part of the nation-building process⁶⁴. After the conflict ended, they were left vulnerable to retribution in Algeria, forcing them to seek refuge in France. However, upon arrival, the Harkis were marginalized. Their loyalty to France was met not with gratitude but with neglect and ostracization, and they struggled to integrate into French society. They were confined to isolated, substandard camps, living there for “weeks or months” while

⁵⁷ Jeannette E. Miller, “A Camp for Foreigners and ‘Aliens’: The Harkis’ Exile at the Rivesaltes Camp (1962–1964),” *French Politics, Culture & Society*, vol. 31, no. 3, 2013, pp. 21–44, <https://www.jstor.org/stable/24517999>.

⁵⁸ “Algerians and French Share Their Stories of the Algerian War,” *FRANCE 24*, March 15, 2022, <https://webdoc.france24.com/algeria-france-war-testimonials/>.

⁵⁹ Algerians who fought alongside the French army during the Algerian War of Independence, often viewed as traitors by Algerian nationalists and marginalized in post-war France.

⁶⁰ “Algerians and French Share Their Stories of the Algerian War,” *FRANCE 24*, March 15, 2022, <https://webdoc.france24.com/algeria-france-war-testimonials/>.

⁶¹ “Jacques Massu, 94; Former French General Regretted Use of Torture,” *Los Angeles Times*, October 29, 2002, <https://www.latimes.com/archives/la-xpm-2002-oct-29-me-passings29.3-story.html>.

⁶² “Algerian Immigration into France | Musée de l’Histoire de l’Immigration,” www.histoire-immigration.fr, n.d., <https://www.histoire-immigration.fr/en/migration-characteristics-by-country-of-origin/algerian-immigration-into-france>.

⁶³ Laura Jeanne Sims, “Rethinking France’s ‘Memory Wars’: Harki Collective Memories, 2003–2010,” *French Politics, Culture & Society*, vol. 34, no. 3, 2016, pp. 83–104, <http://www.jstor.org/stable/44504239>.

⁶⁴ Benjamin Stora, *La Gangrène et l’Oubli: La mémoire de la guerre d’Algérie* (Paris: La Découverte, 1991), p. 8.

some remained “over a decade”⁶⁵. Technically, the Harkis were not prisoners, but they were treated and viewed as such. During their time at the camps, Harkis endured unhealthy living conditions, segregation from the rest of society, military oversight, and “reeducation” classes. These camps effectively “fostered an exilic existence” and “socially excluded them from French society”. The government justified this treatment by claiming that re-education would help Harkis adapt to life in France by giving them “a more profound knowledge of the French language and of [French] morals”⁶⁶. However, this “reeducation” resembled a thinly veiled attempt to impose assimilation while keeping them at arm’s length from mainstream French society. The government’s justification for this treatment underscores a troubling hypocrisy, where the promise of integration was coupled with actions that deepened their exclusion and alienation. This treatment starkly contradicts the very values of liberty, equality, and fraternity that France claims to champion in its national motto, highlighting the limitations of French universalism. By perpetuating a system of exclusion that denied the Harkis their rightful place in the society they had served, it becomes evident that France’s professed commitment to universal principles has often been selective and exclusionary rather than truly inclusive.

Despite facing significant adversity, the Harkis persevered in their efforts to integrate into French society. This persistent admiration for France, despite the discrimination they faced, reflects the effectiveness of the pro-French propaganda that colonial Algerians were subjected to. As one Harki child explained: “It’s true that France locked us up in Harki camps, we did not go to school, we were discriminated against everywhere, we were treated as Arabs, but it’s France. It’s like a mother who hits her children, but she’s nevertheless our mother and we must pardon her and love her”⁶⁷. This troubling statement reflects the deeply ingrained colonial psyche, where the colonized internalize their subjugation and view the colonial power as both oppressors and protectors. The analogy of France as a violent yet beloved mother underscores the psychological dependency created by colonial systems, where loyalty to the colonial power persists despite systemic abuse. It reveals the extent to which colonial structures distort relationships and identities, fostering a sense of obligation to the very system that marginalizes and dehumanizes its subjects. This dynamic highlights the enduring emotional and cultural scars of colonialism, which persist long after political decolonization. Using various websites, the Harkis also diffused an anti-colonial narrative to describe the relationship between France and Algeria. Such examples include statements like “Algeria was a French creation”⁶⁸ and “the falsification of history resides precisely in considering Algeria as a sovereign and constituted power that was

⁶⁵Jeannette E. Miller, “A Camp for Foreigners and ‘Aliens’: The Harkis’ Exile at the Rivesaltes Camp (1962–1964),” *French Politics, Culture & Society*, vol. 31, no. 3, 2013, pp. 21–44, <https://www.jstor.org/stable/24517999>.

⁶⁶ Archives nationales (hereafter AN), 5 AG 1/22, Minister of the Interior, “Rapport sur l’application de la loi du 26 décembre 1961 transmis au Secrétaire général de la Présidence de la République,” p. 104, July 26, 1965.

⁶⁷ Zalamit, post in response to “FIERE DE L’ETRE FILS DE HARKIS,” *Harkis.info*, March 8, 2006, <http://www.harkis.info/portail/viewtopic.php?topic=174&forum=5>.

⁶⁸ Massi, post in response to “[Résolu] - kerchouche Delila,” *Harkis.info*, May 15, 2006, <http://www.harkis.info/portail/viewtopic.php?topic=211&forum=5>.

invaded and occupied by a colonial power”⁶⁹. This mentality explained their actions: “That is why the Harkis fought for France. It was their country. It’s normal to fight for your country against those we call ‘terrorists.’”⁷⁰ This statement highlights a paradox within the Harkis’ identity: while their rhetoric denies Algeria’s historical sovereignty, it also reflects a need to rationalize their alignment with France. By framing their allegiance as “normal” and positioning Algerian nationalists as “terrorists,” they inadvertently adopted colonial narratives that justified French rule. This internalization of colonial ideology not only shaped their perspective but also served as a defense mechanism, enabling them to reconcile their marginalized status within both Algeria and France.

Shockingly, this attitude was also promoted by the French government: on the 23rd of February, 2005, Law No. 2005-158 was passed by the National Assembly, stating that schools should “acknowledge in particular the positive role of the French presence overseas, particularly in North Africa, and accord to the history and sacrifices of combatants in the French army originating from these territories the important place to which they have a right”⁷¹. This is a clear example of how government legislation, ordinary media, and popular culture can directly influence the construction of historical narratives to serve a particular political agenda. By presenting colonialism as a benevolent and constructive force, the atrocities committed during the war by the French are downplayed and legitimized. This also indicates a desire to “misinform” children of the next generation, making them unaware of the country’s shady past, effectively erasing and changing memory. France avoids a critical confrontation with the darker aspects of its colonial past, instead framing its narrative in ways that obscure the suffering and resistance of the Algerian people. This approach shifts the focus toward a sanitized version of history that highlights colonial “contributions” while sidelining its violent realities. By selectively amplifying pro-colonial ideologies, the state creates a narrative that hinders the process of genuine historical reckoning and reconciliation, perpetuating a distorted collective memory.

The law sparked major controversy among the population, and eventually, Chirac repealed it⁷², asserting that “writing history is the business of historians: it should not be circumscribed by laws”⁷³. Chirac was also appalled when he found out that torture was used

⁶⁹ Khader, post in response to “Le film ‘Harkis’ avec Smaïn diffusé sur France 2,” *Coalition nationale des Harkis*, October 17, 2006,

http://www.coalition-harkis.com/component/option.com_fireboard/Itemid,194/func.view/id,272/catid,4/.

⁷⁰ Zinc, post in response to “Un Crime d’État,” *Le Blog de Harkis*, March 25, 2008, <http://harkis.skyrock.com/343398314-Un-crime-d-Etat.html>.

⁷¹ Marcus Otto, “The Challenge of Decolonization: School History Textbooks as Media and Objects of the Postcolonial Politics of Memory in France Since the 1960s,” *Journal of Educational Media, Memory & Society*, vol. 5, no. 1, 2013, pp. 14–32, <https://www.jstor.org/stable/43049652>.

⁷² Itay Lotem, “A Decade After the Riots, France Has Rewritten Its Colonial History,” *The Conversation*, January 25, 2016, <https://theconversation.com/a-decade-after-the-riots-france-has-rewritten-its-colonial-history-53411>.

⁷³ “History Should Not Be Written by Law” says Jacques Chirac (Ce n’est pas à la loi d’écrire l’histoire), quoted by *RFI*, December 11, 2005.

during the war, and when he was informed⁷⁴ that former general Paul Aussaresses claimed torture was “efficient” and that his conscience was clear⁷⁵, he ordered Aussaresses to be suspended from the Legion of Honor and asked Defence Minister Alain Richard to propose eventual disciplinary action⁷⁶. Chirac also tasked historians to quickly shed light on the war period, urging them to study archives made available for the first time last month⁷⁷. Finally, Chirac also made efforts to acknowledge the suffering and sacrifices of the Harkis⁷⁸, such as organizing public ceremonies like the 25 September 2001 Day of National Recognition for the Harkis.

Chirac’s behavior, reflecting a commitment to confronting uncomfortable truths and promoting historical accuracy, was not shared by his successor Sarkozy. The latter refused to apologize for colonial misdeeds, vaguely stating that terrible crimes were committed on both sides and did not elaborate further⁷⁹. He also argued that leaders should focus on the future and not “beat their breasts”⁸⁰, maintaining that building a constructive relationship required looking forward rather than dwelling on the past—a stance that echoed the selective memory politics of the Vichy Syndrome debate, where the focus on unity often overshadowed the darker truths of collaboration⁸¹. When around 160 Algerian and Moroccan politicians, lawyers, and human rights advocates published a joint appeal for France to acknowledge the “trauma caused by the colonization of Algeria,” as an effort to heal ties, it was ignored⁸². When Sarkozy planned to visit Algeria, he was not welcomed with open arms. Mohamed Said Abadou, an influential Algerian war veterans’ leader was quoted saying the president was not welcome until he apologized for the colonial past: “Sarkozy is not welcome in Algeria [...] and we won’t turn the page with France until we get an apology.”⁸³ To top it all off, Sarkozy’s previous “Kärcher” comments, in which he referred to using a high-pressure cleaner to rid the Paris suburbs of “scum” during the 2005 riots⁸⁴, further inflamed tensions and reinforced the perception that he held deeply negative

⁷⁴ Human Rights Watch, “Letter to French President Jacques Chirac Calling for War Crimes Investigation,” May 14, 2001, <https://www.hrw.org/news/2001/05/13/letter-french-president-jacques-chirac-calling-war-crimes-investigation>.

⁷⁵ “Chirac ‘Horried’ by War Claims,” *CNN.com*, May 4, 2001, <https://edition.cnn.com/2001/WORLD/europe/05/04/france.algeria/>.

⁷⁶ Barry James and *International Herald Tribune*, “General’s Confessions of Torture Stun France,” *The New York Times*, May 5, 2001, <https://www.nytimes.com/2001/05/05/news/generals-confessions-of-torture-stun-france.html>.

⁷⁷ “Chirac ‘Horried’ by War Claims,” *CNN.com*, May 4, 2001, <https://edition.cnn.com/2001/WORLD/europe/05/04/france.algeria/>.

⁷⁸ “Chirac Hails Algerians Who Fought for France,” *The Telegraph*, September 26, 2001.

⁷⁹ Agence France-Presse, “In Algeria, Sarkozy Denounces Colonialism,” *The New York Times*, December 4, 2007, <https://www.nytimes.com/2007/12/04/world/africa/04algeria.html>.

⁸⁰ Reuters, “Sarkozy Tells Algeria: No Apology for the Past,” *Reuters*, August 10, 2007, <https://www.reuters.com/article/world/sarkozy-tells-algeria-no-apology-for-the-past-idUSL10638737/>.

⁸¹ L. Aaron, “Nicolas Sarkozy’s Conservative Revolution: The Price for France,” *The Asia-Pacific Journal: Japan Focus*, August 2007.

⁸² Reuters, “Sarkozy Tells Algeria: No Apology for the Past,” *Reuters*, August 10, 2007, <https://www.reuters.com/article/world/sarkozy-tells-algeria-no-apology-for-the-past-idUSL10638737/>.

⁸³ News24, “Sarkozy ‘Owes Algeria Apology,’” *News24*, November 29, 2007, <https://www.news24.com/news24/sarkozy-owes-algeria-apology-20071129>.

⁸⁴ “‘Nettoyer Au Kärcher’: How High-Pressure Cleaning Became a Political Issue in France.” *Galaxus*, May 22, 2022.

views toward communities of North African descent, including Algerians. His remarks were inflammatory and derogatory, solidifying harmful stereotypes that cast Algerian and other North African immigrants as violent and unassimilable. The rift between these communities and the French state was aggravated, making reconciliation and integration increasingly difficult.

When François Hollande came into office in 2012, succeeding Sarkozy, the dominant narrative was not altered. He did not formally apologize to the Algerians, “[stopping] short of apologising for the past,” but instead acknowledged France’s colonial legacy by stating: “I recognize the suffering the colonial system has inflicted.”⁸⁵ However, while previous presidents remained silent about the October Massacre, during which French police violently suppressed peaceful Algerian protesters in Paris, resulting in dozens of deaths, Hollande stood out by addressing it; he was “the first high-ranking official to own up to the tragedy, unnerving the French right. Two months later, he doubled down with his apologetic speech to the Algerian Parliament”⁸⁶, a stark contrast from the silence maintained by previous presidents about the 1961 massacre of peaceful Algerian protesters. Yet, it is important to remember that there was still no apology, only an acknowledgment.

Additionally, similarly to Sarkozy, Hollande advocated for turning the page and embarking on a fresh start, asserting that this new beginning must be “supported by a base,” which he defined as the truth: “nothing is built in secretiveness, forgetting, denial.”⁸⁷ Yet Hollande's reluctance to offer a formal apology suggests that his acknowledgment of past traumas might have been more about creating a reason to move on than a genuine effort to address them. This reflects how memory, in such cases, often becomes utilitarian—a tool employed to serve political and social objectives rather than a sincere reckoning with history. In this instance, memory promotes closure or appeases public sentiment without requiring meaningful accountability or systemic change. The question, then, is not merely about memory itself but how it is wielded, often as an instrument to justify selective narratives or strategic forgetting. While Hollande did recognize the suffering inflicted by the colonial system, his hesitation to provide a complete apology raises questions about whether mere acknowledgment is enough to overcome a legacy of secrecy and denial. Indeed, Professor of Middle East Studies, Camille Pecastaing comments on the long-term effectiveness of this apology, claiming it remains questionable: “it is not clear that, aside from their value as theater, Hollande’s regrets will have measurable effect, one way or the other.” Pecastaing compares his gesture, seen as a one-off statement, to President Obama’s 2009 Cairo speech, which ultimately did little to change the dynamics of Middle Eastern relations. Hollande’s actions were viewed as unlikely to bridge the

<https://www.galaxus.de/en/page/nettoyer-au-karcher-how-high-pressure-cleaning-became-a-political-issue-in-france-23763>.

⁸⁵ “François Hollande Acknowledges Algerian Suffering under French Rule,” *The Guardian*, December 20, 2012, <https://www.theguardian.com/world/2012/dec/20/francois-hollande-algerian-suffering-french>.

⁸⁶ Camille Pecastaing, “THE POLITICS OF APOLOGY: Hollande and Algeria,” *World Affairs*, vol. 175, no. 6, 2013, pp. 51–56, <http://www.jstor.org/stable/43556163>.

⁸⁷ “François Hollande Acknowledges Algerian Suffering under French Rule,” *The Guardian*, December 20, 2012, <https://www.theguardian.com/world/2012/dec/20/francois-hollande-algerian-suffering-french>.

rift between France and Algeria, with Pecastaing suggesting that historical grievances may not significantly influence current relations: “time passes and the crimes of French colonialism do not reflect on contemporary French society.”⁸⁸ Acknowledgment alone is insufficient. Without a formal apology, France’s efforts to address its colonial past cannot be considered a genuine attempt to turn the page. Moving forward requires more than just recognition; it demands a sincere apology.

Years later, with Emmanuel Macron running for the presidency, there was still no apology. During his 2017 election campaign, Macron condemned the occupation of Algeria as a “crime against humanity” and described French actions as “genuinely barbaric.”⁸⁹ However, despite this strong denunciation, he did not actually apologize, explicitly stating that there would be “no repentance nor apologies.”⁹⁰ While denouncing past actions is a step forward, it is clear that a genuine apology is necessary to address historical grievances and move toward real reconciliation. There has already been lots of condemnation; more needs to be done now, and the crucial step forward is offering formal apologies.

This necessity for deeper accountability continues to be highlighted, this time by recent revelations. In a significant admission, President Macron publicly acknowledged for the first time that French soldiers were responsible for the torture and murder of a prominent Algerian independence figure and subsequently covered up his death⁹¹. French authorities “previously claimed that he had committed suicide while in detention, a lie that his widow and other family members had campaigned for years to see overturned.”⁹² Macron’s reluctance to discuss the issue until he was pressured suggests an underlying tendency to conceal such historical truths. This raises a critical question: how many other stories remain untold, obscured either by historical neglect or by the lack of sufficient public pressure to compel political leaders to confront and address them?

Conclusion

The Nazi Occupation and the Algerian War leave France with deep dilemmas and they reveal the complicated interplay between memory and identity and justice. Some argue that silencing parts of memory is necessary to encourage solidarity and ease nation-building but this approach risks creating a fragile foundation that is built on denial and omission. France actively

⁸⁸ Camille Pecastaing, “The Politics of Apology: Hollande and Algeria,” *World Affairs*, vol. 175, no. 6, 2013, pp. 51–56, <http://www.jstor.org/stable/43556163>.

⁸⁹ “Macron Calls Colonialism a ‘Grave Mistake’ during Visit to Ivory Coast.” *France 24*, December 22, 2019, <https://www.france24.com/en/20191222-france-president-macron-on-official-visit-to-ivory-coast-calls-colonialism-a-grave-mistake>.

⁹⁰ “Macron Calls Colonialism a ‘Grave Mistake’ during Visit to Ivory Coast.” *France 24*, December 22, 2019, <https://www.france24.com/en/20191222-france-president-macron-on-official-visit-to-ivory-coast-calls-colonialism-a-grave-mistake>.

⁹¹ “Ali Boumendjel: France Admits ‘Torture and Murder’ of Algerian Nationalist,” *BBC News*, March 3, 2021, <https://www.bbc.com/news/world-europe-56262963>.

⁹² *Al Jazeera*, “Macron Admits France Murdered Top Algerian Independence Figure,” March 3, 2021, <https://www.aljazeera.com/news/2021/3/3/macron-admits-france-murdered-top-algerian-independence-figure>.

engages with its past to show how it can wield memory in order to craft narratives that zero in on national cohesion and also overlook historical accountability.

To achieve true unity, we must address and heal from suppressed trauma and embrace all narratives rather than selectively choosing only some. An engaged and truthful reckoning demands a multidimensional approach that links acknowledgment and reparative justice. We must honor the victims of both the Nazi Occupation and colonial violence and integrate marginalized histories through educational reforms and provide reparations. Reparations are a way to acknowledge the harm done to affected communities actively and they also represent a commitment to rebuilding trust with those same communities. Germany's reparations to Holocaust survivors are shown to depict how such actions can contribute to reconciliation and address systemic injustices.

To foster an inclusive collective memory, public spaces for dialogue must be created where competing narratives can coexist and be critically examined. Initiatives like truth commissions, testimonial projects, or state-sponsored efforts can provide platforms for acknowledging silenced voices and confronting the complexities of both colonial and wartime histories. By actively engaging with the past rather than selectively forgetting it, France can dismantle the selective narratives that have long shaped its understanding of itself and foster a collective identity that embraces accountability and inclusivity.

Ultimately, embracing the truth must involve using apology as a starting point for change, not as a conclusion to close the doors of the past. Reparations, education, and inclusive memorialization should work together to confront historical injustices and honor the diversity of experiences that constitute France's national story. This is not about diminishing the country's achievements but enriching its identity by confronting contradictions—acknowledging that strength lies in vulnerability and greatness in humility. Such efforts are not only essential for historical reconciliation but are also vital for shaping a Republic that truly lives up to its principles of liberty, equality, and fraternity. Only by listening to the silenced, honoring the unspoken truths, and committing to transparency and accountability can France build a future grounded in justice and the shared memories of all its citizens. This journey demands moral courage and promises a national identity that celebrates its diversity rather than erasing it.

Works Cited

- Aaron, L. "Nicolas Sarkozy's Conservative Revolution: The Price for France." *The Asia-Pacific Journal: Japan Focus*, August 2007.
- Agence France-Presse. "In Algeria, Sarkozy Denounces Colonialism." *The New York Times*, December 4, 2007, <https://www.nytimes.com/2007/12/04/world/africa/04algeria.html>.
- Al Jazeera. "Macron Admits France Murdered Top Algerian Independence Figure." *Al Jazeera*, March 3, 2021, <https://www.aljazeera.com/news/2021/3/3/macron-admits-france-murdered-top-algerian-independence-figure>.
- "Ali Boumendjel: France Admits 'Torture and Murder' of Algerian Nationalist." *BBC News*, March 3, 2021, <https://www.bbc.com/news/world-europe-56262963>.
- "Algerian Immigration into France." *Musée de l'Histoire de l'Immigration*, n.d., <https://www.histoire-immigration.fr/en/migration-characteristics-by-country-of-origin/algerian-immigration-into-france>.
- "Algerians and French Share Their Stories of the Algerian War." *FRANCE 24*, March 15, 2022, <https://webdoc.france24.com/algeria-france-war-testimonials/>.
- Archives nationales. Rapport sur l'application de la loi du 26 décembre 1961 transmis au Secrétaire général de la Présidence de la République. 26 July 1965, p. 104. Archives nationales, 5 AG 1/22, Minister of the Interior.
- Bernardot, Marc. *Camps d'étrangers* (in French). Paris: Terra, 2008.
- Bergen, Bernard J. *The Banality of Evil*. Lanham, MD: Rowman & Littlefield Publishers, 2000.
- Bracher, Nathan. "Bruckner and the Politics of Memory: Repentance and Resistance in Contemporary France." *South Central Review*, vol. 24, no. 2, 2007, pp. 54–70, <https://www.jstor.org/stable/40039981>.
- Carrier, Peter. *Holocaust Monuments and National Memory Cultures in France and Germany Since 1989*. New York: Berghahn Books, 2005.
- "Chirac Admits France's Complicity with Nazis: Europe: For the First Time, a French President Acknowledges the Nation's Role in the Deportation of Jews during World War II." *Los Angeles Times*, July 17, 1995, <https://www.latimes.com/archives/la-xpm-1995-07-17-mn-24867-story.html>.
- "Chirac Hails Algerians Who Fought for France." *The Telegraph*, September 26, 2001.
- "CNN.com - Chirac 'Horrificed' by War Claims - May 4, 2001." *CNN.com*, May 4, 2001, <https://edition.cnn.com/2001/WORLD/europe/05/04/france.algeria/>.
- Jean-Paul Cointet. *Épuration légale: 400 000 dossiers, moins de 800 morts*. *Historia* (in French).
- Davis, Mark. "How World War II Shaped Modern France." *Euronews*, May 5, 2015, <https://www.euronews.com/2015/05/05/how-world-war-ii-shaped-modern-france>.
- Elkabbach, Jean-Pierre. *François Mitterrand, Conversation avec un Président*. Directed by Jean-Pierre Elkabbach. 1994. Documentary.

- France 24. "France Remembers the Algerian War, 50 Years On." *France 24*, March 16, 2012, <https://www.france24.com/en/20120316-commemorations-mark-end-algerian-war-independence-france-evian-accords>.
- "François Hollande Acknowledges Algerian Suffering under French Rule." *The Guardian*, December 20, 2012, <https://www.theguardian.com/world/2012/dec/20/francois-hollande-algerian-suffering-french>.
- "FRANÇOIS MITTERRAND Répondra Aux Questions de Jean-Pierre Elkabbach." *L'Humanité*, September 11, 1994, <https://www.humanite.fr/-/-/francois-mitterrand-repondra-aux-questions-de-jean-pierre-elkabbac>.
- "History Should Not Be Written by Law" says Jacques Chirac (Ce n'est pas à la loi d'écrire l'histoire). Quoted by *RFI*, December 11, 2005.
- Human Rights Watch. "Letter to French President Jacques Chirac Calling for War Crimes Investigation." May 14, 2001, <https://www.hrw.org/news/2001/05/13/letter-french-president-jacques-chirac-calling-war-crimes-investigation>.
- Jackson, Julian. *De Gaulle*. Cambridge, MA: Harvard University Press, 2018.
- James, Barry, and International Herald Tribune. "General's Confessions of Torture Stun France." *The New York Times*, May 5, 2001, <https://www.nytimes.com/2001/05/05/news/generals-confessions-of-torture-stun-france.html>.
- Jaspers, Karl, and E. B. Ashton. *The Question of German Guilt*. New York: Fordham University Press, 1965, <https://doi.org/10.2307/j.ctt13wzz9w>.
- "Jacques Massu, 94; Former French General Regretted Use of Torture." *Los Angeles Times*, 29 Oct. 2002, <https://www.latimes.com/archives/la-xpm-2002-oct-29-me-passings29.3-story.html>.
- Khader. "Post in Response to 'Le Film "Harkis" avec Smaïn Diffusé sur France 2.'" *Coalition nationale des Harkis*, 17 Oct. 2006, http://www.coalition-harkis.com/component/option,com_fireboard/Itemid,194/func,view/id,272/catid,4/.
- Kyte, George W. "War Damage and Problems of Reconstruction in France, 1940-1945." *Pacific Historical Review*, vol. 15, no. 4, 1946, pp. 417–26, <https://doi.org/10.2307/3635778>.
- La Décennie Mitterrand. Paris: Éditions du Seuil, coll. "Points," 2001.
- Lotem, Itay. "A Decade After the Riots, France Has Rewritten Its Colonial History." *The Conversation*, January 25, 2016, <https://theconversation.com/a-decade-after-the-riots-france-has-rewritten-its-colonial-history-53411>.
- Massi. "Post in Response to '[Résolu] - kerchouche Delila.'" *Harkis.info*, 15 May 2006, <http://www.harkis.info/portail/viewtopic.php?topic=211&forum=5>.
- Miller, Jeannette E. "A Camp for Foreigners and 'Aliens': The Harkis' Exile at the Rivesaltes Camp (1962–1964)." *French Politics, Culture & Society*, vol. 31, no. 3, 2013, pp. 21–44, <https://www.jstor.org/stable/24517999>.
- "'Nettoyer Au Karcher': How High-Pressure Cleaning Became a Political Issue in France." *Galaxus*, May 22, 2022,

- <https://www.galaxus.de/en/page/nettoyer-au-karcher-how-high-pressure-cleaning-became-a-political-issue-in-france-23763>.
- News24. "Sarkozy 'Owes Algeria Apology.'" News24, November 29, 2007, <https://www.news24.com/news24/sarkozy-owes-algeria-apology-20071129>.
- "Ombres et lumières de la révolution algérienne." *Le Monde diplomatique*, November 1, 1982.
- Otto, Marcus. "The Challenge of Decolonization: School History Textbooks as Media and Objects of the Postcolonial Politics of Memory in France Since the 1960s." *Journal of Educational Media, Memory & Society*, vol. 5, no. 1, 2013, pp. 14–32, <https://www.jstor.org/stable/43049652>.
- Pecastaing, Camille. "The Politics of Apology: Hollande and Algeria." *World Affairs*, vol. 175, no. 6, 2013, pp. 51–56, <https://www.jstor.org/stable/43556163>.
- Péan, Pierre. *Une Jeunesse Française*. Paris: Fayard, 2014.
- "Pourquoi la rafle n'a pas atteint son objectif." AIDH.org, p. 52, <http://aidh.org/pdf/rafle-obstacle>.
- Reuters. "Sarkozy Tells Algeria: No Apology for the Past." Reuters, August 10, 2007, <https://www.reuters.com/article/world/sarkozy-tells-algeria-no-apology-for-the-past-idUSL10638737/>.
- Scullion, Rosemarie. "Unforgettable: History, Memory, and the Vichy Syndrome." *Studies in 20th & 21st Century Literature*, vol. 23, no. 1, 1999, <https://doi.org/10.4148/2334-4415.1452>.
- Shillington, Kevin. *Encyclopedia of African History: 3-Volume Set*. New York: Routledge, 2013.
- Simons, Marlise. "Chirac Affirms France's Guilt in Fate of Jews." *The New York Times*, July 17, 1995, <https://www.nytimes.com/1995/07/17/world/chirac-affirms-france-s-guilt-in-fate-of-jews.html>.
- Sims, Laura Jeanne. "Rethinking France's 'Memory Wars': Harki Collective Memories, 2003–2010." *French Politics, Culture & Society*, vol. 34, no. 3, 2016, pp. 83–104, <http://www.jstor.org/stable/44504239>.
- Stora, B. *La gangrène et l'oubli: La mémoire de la guerre d'Algérie*. Paris: La Découverte, 1991.
- Vidal-Naquet, Pierre. *Les crimes de l'armée française: Algérie, 1954-1962*. Paris: La Découverte, 2001.
- Vie-publique.fr. Archived April 12, 2008, at the Wayback Machine. Extracts from the interview François Mitterrand gave to Radio J on November 13, 1992.
- Zalamit. Post in response to "FIERE DE L'ETRE FILS DE HARKIS." Harkis.info, March 8, 2006, <http://www.harkis.info/portail/viewtopic.php?topic=174&forum=5>.
- Zinc. "Post in Response to 'Un Crime d'État.'" *Le Blog de Harkis*, 25 Mar. 2008, <http://harkis.skyrock.com/343398314-Un-crime-d-Etat.html>.
- Figure 1: The Deportation of Jews from France, held by the World Holocaust Remembrance Center. The image shows "German and French policemen, conducting the deportation of Jews from the old harbor district in Marseilles, January 1943." Yad Vashem, <https://www.yadvashem.org/holocaust/france/deportation-from-france.html>

The Effects of Physical Activity and Sleep on Stress and Happiness: A Comparative Study of General and International Alternative School Students By Yoseop Jeong

Abstract

This study aimed to analyze the correlations between stress, physical activity, and sleep among students from general schools and international alternative schools in South Korea, comparing the differences to derive practical implications for reducing adolescent stress. Data were collected using the *Youth Health Behavior Survey* for general schools and a tailored survey for international alternative schools, with responses analyzed using IBM SPSS Statistics 22. Results revealed that while international alternative school students reported higher levels of academic and appearance-related stress, they also exhibited higher subjective happiness, greater physical activity participation, and better sleep satisfaction compared to general school students. The findings highlight the need for balanced academic expectations, tailored stress management programs, and enhanced physical activity and sleep initiatives to address adolescent stress effectively.

Introduction

In recent years, youth mortality rates and life satisfaction levels have emerged as pressing and crucial social issues in South Korea. The *2024 Youth Statistics*, jointly published by the Ministry of Gender Equality and Family and the National Youth Policy Institute in late May, include intense alarming findings that underscore these concerns (Ministry of Gender Equality and Family, 2024). This report comprehensively examines the physical and mental conditions of adolescents aged 9 to 24, focusing on their level of education, health, leisure, and social participation. Particularly troubling is the fact that suicide has remained the leading cause of death among adolescents for the past 12 years, highlighting the urgent need for effective and long-term solutions.

At the core of these issues lies stress, a fundamental driver of adolescent suicides. Accumulated stress can lead to suicidal impulses, ultimately resulting in tragic outcomes. Between 2017 and 2022, the suicide rate among male and female adolescents has shown an overall increasing trend (Statistics Korea, 2023). According to the 2023 Cause of Death Statistics, the total number of suicides in 2023 reached 13,978, an increase of 1,072 cases (8.3%) compared to 2022, with a suicide mortality rate of 27.3 per 100,000 people, marking an 8.5% rise (Ministry of Health and Welfare, 2024). In a society where suicide rates are steadily climbing, adolescent stress, particularly as it relates to suicide, is recognized as a critical issue.

Many factors, such as low self-esteem, hopelessness, daily stress, negative family experiences, school-related stress, and peer relationships, contribute to adolescent suicide. However, stress stands out as the overarching cause (Shin et al., "Factors Influencing Depression, Suicidal Ideation, and Attempts among Adolescents," *School Social Work*, vol. 27, 2014, pp. 25-50). Stress and depression interact as significant factors triggering suicidal impulses. In 2022, the perceived stress rate among middle and high school students in Korea was 41.3%, an

increase of 2.5% from 2021. By school level, the rates were 43.0% for high school students and 39.8% for middle school students (Ministry of Gender Equality and Family, 2023). These statistics illustrate the urgent need to explore practical solutions to alleviate stress among adolescents who endure psychological challenges annually.

Amid efforts to address adolescent stress, research suggests that physical activity and sufficient sleep can significantly reduce stress levels (Seo et al., 2022). Physical activity during adolescence enhances social bonds and serves as a constructive outlet for stress relief. Additionally, adequate sleep is vital; sleep deprivation can lead to growth delays, fatigue, headaches, and impaired cognitive functions such as memory and concentration, thereby posing critical challenges for adolescents (Yoon et al., 2007).

Adolescent stress frequently stems from South Korea's intense academic competition. Researchers have explored various alternative schooling methods to address structural issues in education. Among these, international alternative schools, which combine the American educational system with a Christian worldview, have gained recognition for overcoming the limitations of conventional public education. These schools aim to foster holistic development and impart practical wisdom for life, yielding high levels of satisfaction among their students. Consequently, comparing stress levels between students in general high schools and international alternative schools offers valuable insights.

Korean students' school environments significantly influence their stress and life satisfaction in modern society. International alternative schools and general schools differ substantially in terms of teaching methods, educational philosophies, and student composition, all of which affect students' mental and physical well-being. Therefore, this study aims to compare the stress levels of students from these two types of schools and examine their correlations with physical activity and sleep. By analyzing the causes of adolescent stress and comprehensively comparing its relationship with physical activity and sleep, this research seeks to provide a new perspective on promoting adolescent health.

Background

<About International Alternative Schools>

Before delving into the definition of alternative schools, it is essential to first understand their core nature. Shim Serin (2023) says that alternative schools are places that try to fix the problems with regular public schools by using different ways of teaching that are based on each student's needs and goals, without having to follow set rules for entrance exams. The limitations of traditional public education often refer to the excessive competition and overwhelming academic workload. Under the immense influence of parental expectations and societal pressure for prestigious university admissions, South Korean adolescents tend to perceive college acceptance and material success as ultimate goals. However, they are rarely given opportunities to explore what they genuinely enjoy and excel at. Alternative schools were established to

address these structural limitations of public education and to broaden the development of students' diverse talents and potential.

The essence of alternative schools lies not merely in the transmission of academic knowledge based on textbooks but in nurturing the skills and experiences necessary for life. Moving away from rote learning methods and rigidly structured daily routines, alternative schools focus on fostering students' autonomy and creativity while offering lessons that emphasize values essential to future society. Their dynamic and diversified curricula are designed to reflect individual students' needs and interests, exerting a positive influence on adolescents in South Korea. Alternative schools are steadily positioning themselves as institutions that complement the shortcomings of conventional education systems while providing students with broader perspectives and diverse experiences (Jung Yoosung, 1997).

As the scope of activities in alternative schools expands and the number of students continues to grow, various alternative schools with unique philosophies and objectives have emerged. Among these, international alternative schools stand out as a type of alternative education system that combines international education with character development to cultivate global talent. Emerging in the early 2000s, these schools have steadily increased in number (Nam Jinseok, 2009). International alternative schools focus on improving students' global competencies, multicultural understanding, and language skills to prepare them for international competition and global leadership (Kang Daeyong, 2018). General alternative schools focus on holistic growth, creativity, and community spirit through a flexible and self-directed learning environment.

International alternative schools, aimed at fostering international competitiveness and global leadership, exhibit distinct characteristics that set them apart from public education. First, their curricula align with high-level international standards such as IB (International Baccalaureate) and AP (Advanced Placement), enabling students to build academic competencies competitive with global peers. Additionally, unlike general schools, international alternative schools typically have smaller class sizes, which encourages closer relationships and collaboration among students (Seo Dami, 2017). Students naturally bond with peers, as well as senior and junior students, creating a comfortable and unique school atmosphere. Juniors can seek practical academic and college admission advice from seniors, who, in turn, actively communicate and maintain close relationships with their juniors. Moreover, the relationship between teachers and students is further strengthened through continuous communication and various school events. Activities such as bazaars, sports days, and retreats foster a sense of unity among students and teachers, enhancing their mutual bonds over time.

Another core value of international alternative schools is their focus on ethical education based on Christian values, emphasizing moral leadership and community spirit (Ham Youngjoo, 2015). Through this, students with a foundation in Christian values significantly improve their problem-solving and communication skills. These distinctions from public education contribute substantially to nurturing students in international alternative schools as globally competitive individuals.

Thus, international alternative schools not only overcome the rigid and competitive limitations of traditional entrance exam-focused education but also respect students' individuality and character, promoting creativity and moral development. Furthermore, the friendly relationships between peers and between students and teachers lead to high levels of satisfaction with school life and excellent academic outcomes (Seo Dami, 2017).

<Adolescent Stress>

Adolescence is a period characterized by rapid physical and emotional changes. During this time, adolescents acquire and experience the social relationships, moral values, and knowledge necessary to grow and develop into healthy members of society. These experiences allow them to adapt to various social demands (Kang Sun-Young, 2010). As self-identity forms and social perspectives are shaped; adolescents build self-esteem by interacting with diverse individuals. In schools, which function as small-scale societies, they encounter various challenges and experiences that can have both positive and negative effects depending on the circumstances. Positive school life adaptation can significantly impact not only adolescence but also the quality of life in adulthood. Conversely, stress and maladjustment in school life are known to lead to mental, psychological, and physical health issues as well as problem behaviors such as truancy, running away from home, and delinquency (Kim Shin-Ae, Lee Hyung-Sil, and Lim Soo-Kyung, 2008; Koo Mi-Hyang, 2012).

Stress during this critical developmental period cannot be overlooked. According to the 2023 Adolescent Health Behavior Survey, 30.8% of male high school students and 44.2% of female high school students reported experiencing stress, with female students recording a relatively higher rate. The primary sources of stress were academic performance, career concerns, schoolwork, appearance, and conflicts with parents. In 2015, the Korea Institute for Health and Social Affairs measured the academic stress index of South Korean adolescents by applying UNICEF's *2013 Report on the Subjective Well-being of Adolescents in Affluent Countries*. The findings revealed that South Korean adolescents had the highest academic stress index among the 29 countries surveyed, at 50.5% (Hwang Soon-Gil, Lee Yun-Joo, and Lee Hye-Jeong, 2017). Various large-scale surveys and research studies have consistently shown that South Korean adolescents experience high levels of stress. Within the realm of perceived stress, adolescents report stress in multiple categories, including academics, grades and exams, college admissions and career paths, peer relationships, teacher relationships, family issues, romantic relationships, appearance, and personality (Hwang Soon-Gil, Lee Yun-Joo, and Lee Hye-Jeong, 2017).

Among these factors, research has identified sleep and physical activity as closely related to stress. Adolescents who experience high stress levels are 1.36 times more likely to suffer from sleep deprivation compared to those with lower stress levels, confirming that stress increases the risk of insufficient sleep. Research also demonstrates that middle and high school students who get more sleep experience less stress and are less likely to consider suicide. This shows that getting enough sleep greatly lowers stress (Korea Disease Control and Prevention Agency,

2021). Meanwhile, Kim et al. (2021) demonstrated that physical activity among adolescents positively and significantly impacts not only stress reduction but also quality of life. This finding aligns with statistics showing that adolescents who do not engage in physical activity have higher stress perception rates compared to those who do (Korea Disease Control and Prevention Agency, 2021).

The type of school adolescents attend also affects the types and levels of stress they experience, as factors such as admission motivations and goals for higher education vary. Public education, with its primary focus on advancing to higher education or securing employment, imposes significant stress on students striving to achieve these goals (Seo Kyung-Hee, 2016). Given the substantial time adolescents spend in school and the burdens of academic performance and peer relationships, school life adaptation is directly and inversely related to stress (Kim Jung-Hyun, Kim Sung-Byuk, and Jung In-Kyung, 2014). To alleviate the stress caused by academic pressure and enhance school life adaptation, it is essential to transform schools from competitive, exam-driven environments into holistic educational spaces. These spaces should foster students' potential, individuality, and healthy development, promoting overall well-being and satisfaction in school life.

Method

<Study Subjects>

This study aims to compare the stress levels of students in South Korea's general schools and international alternative schools and investigate their correlations with physical activity and sleep. For this purpose, data were collected from the *Youth Health Behavior Survey* for general schools and from surveys conducted specifically for international alternative schools.

To analyze data from general schools, the study utilized the 19th *Youth Health Behavior Survey* conducted in 2023 by the Korea Disease Control and Prevention Agency (KDCA). The *Youth Health Behavior Survey* is an anonymous, self-administered online survey targeting middle and high school students (grades 7 to 12) in South Korea. It has been conducted annually since 2005, the national approved statistics (Approval No. 117058) have conducted the survey annually to assess the health behaviors of South Korean adolescents, including smoking, drinking, diet, and physical activity. Since 2005 as nationally approved statistics (Approval No. 117058) to assess the health behaviors of South Korean adolescents, such as smoking, drinking, diet, and physical activity. The purpose of the survey is to generate statistical data on adolescent health behaviors, which serve as critical evidence for planning and evaluating national health policies and programs. The 19th survey included a total of 52,880 participants, making it highly representative of South Korean adolescents' overall health behaviors.

For the international alternative school analysis, the study used Google Forms to replicate selected items from the *Youth Health Behavior Survey* focusing on stress, physical activity, and sleep. This survey was conducted across six international alternative schools in South Korea, located in cities such as Yongin, Jeju, and Seoul, with a total of 73 participants.

<Survey Items and Research Tools>

The collected data were analyzed using IBM SPSS Statistics 22. The analysis process included the following steps: First, general characteristics of participants were derived using frequency analysis. Second, crosstabs analysis was implied to compare levels and causes of stress, as well as physical activity and sleep patterns, by gender and school level regarding middle school and high school. Third, the results of general schools and international alternative schools were then compared.

Result

<General Characteristics>

The general characteristics of the study are summarized as follows: The gender distribution between general schools and international alternative schools was relatively balanced, though international alternative schools showed a higher proportion of male students (60.3%) compared to female students (39.7%). In terms of school level, general schools had a nearly even distribution of middle school (53.7%) and high school students (46.3%). However, international alternative schools showed a slightly higher proportion of middle school students (57.6%) compared to high school students (42.4%).

Regarding academic achievement, the most frequent response in both school types was “average,” with 29.9% in general schools and 25.0% in international alternative schools. However, the proportion of students who responded with “very low” and “low” academic achievement was higher in international alternative schools (44.4%) compared to general schools (31.4%).

Economic status was predominantly reported as “average” in both school types, with 46.6% in general schools and 39.5% in international alternative schools. However, international alternative schools had a higher proportion of students reporting “very high” economic status (14.2%) compared to general schools (11.5%), while the proportion reporting “very low” economic status (8.0%) was also relatively higher in international alternative schools.

Characteristic	General School (n=52,880/100%)	International Alternative School (n=73/100%)
Gender		
Male	26,769 (50.6%)	44 (60.3%)
Female	26,111 (49.4%)	29 (39.7%)
School Level		
Middle School	28,401 (53.7%)	43 (57.6%)
High School	24,479 (46.3%)	30 (42.4%)
Academic Achievement		
Very High	6,935 (13.4%)	8 (11.7%)
High	13,116 (25.3%)	14 (19.0%)
Average	15,484 (29.9%)	18 (25.0%)
Low	11,380 (21.9%)	19 (25.6%)
Very Low	4,933 (9.5%)	14 (18.8%)
Economic Status		
Very High	5,984 (11.5%)	10 (14.2%)
High	15,904 (30.7%)	18 (24.8%)
Average	24,143 (46.6%)	29 (39.5%)
Low	4,807 (9.3%)	10 (13.4%)
Very Low	1,009 (1.9%)	6 (8.0%)

table. 1 : General Characteristics of Research Subjectss

<Subjective Happiness>

In response to the question, "How happy do you generally feel?" the most common answer among students from general schools was "Somewhat happy," at 36.5%. "Very happy" was reported by 21.5% of respondents, while "Somewhat unhappy" and "Very unhappy" had lower response rates of 10.2% and 1.9%, respectively.

Among students from international alternative schools, the most frequent response was also "Somewhat happy," at 34.2%. However, the proportion of students who answered "Very happy" was 27.4%, 5.9 percentage points higher than that of general schools. On the other hand, the response rate for "Very unhappy" was 5.5%, slightly higher than that of general schools.

Subjective Happiness	General School (n=52,880)	International Alternative School (n=73)
Very Happy	11,383 (21.5%)	20 (27.4%)
Somewhat Happy	19,310 (36.5%)	25 (34.2%)
Neutral	15,816 (29.9%)	20 (27.4%)
Somewhat Unhappy	5,383 (10.2%)	4 (5.5%)
Very Unhappy	988 (1.9%)	4 (5.5%)

table. 2 : Results of Subjective Happiness

<Stress Perception Rate>

In response to the question, "How often do you feel stressed in your daily life?" 45.1% of students from general schools reported "a little," while 28.1% answered "a lot," and 9.1% responded with "Very much."

For students in international alternative schools, the most common responses were also "a little" (41.1%) and "a lot" (20.5%). However, the proportion of students who answered "Very much" was higher at 13.7%, compared to students in general schools.

This indicates that while moderate levels of stress are prevalent in both school types, the intensity of stress ("Very much") tends to be more pronounced in international alternative schools.

Stress Awareness Level	General School (n=52,880)	International Alternative School (n=73)
Extremely Stressed	4,836 (9.1%)	10 (13.7%)
Highly Stressed	14,863 (28.1%)	15 (20.5%)
Moderately Stressed	23,874 (45.1%)	30 (41.1%)
Barely Stressed	7,794 (14.7%)	10 (13.7%)
Not Stressed at All	1,513 (2.9%)	8 (11.0%)

table. 3 : Stress Awareness

<Major Causes of Stress>

Students from general schools identified "relationships with peers and seniors and juniors" (34.2%) and "academic performance and career concerns" (24.7%) as the primary causes of their stress.

In contrast, students from international alternative schools overwhelmingly cited "academic performance and career concerns" as their top stressor, with 53.4% selecting this option—a significantly higher proportion compared to general schools. Additionally, stress

related to "appearance" was also notable among international alternative school students, with 15.1% identifying it as a key source of stress.

This suggests that while interpersonal relationships play a significant role in stress levels at general schools, academic pressures are a dominant factor in international alternative schools, reflecting the differing educational priorities and environments.

Stress Cause	General School (n=52,880)	International Alternative School (n=73)
Academic and Career Pressure	12,274 (24.7%)	39 (53.4%)
Appearance	2,131 (4.0%)	11 (15.1%)
Conflict with Parents	4,781 (9.0%)	5 (6.8%)
Peer and Senior Relationships	18,091 (34.2%)	13 (17.8%)
Others	15,603 (28.1%)	5 (6.8%)

table. 4 : Major Causes of Stress

<Days of Physical Activity>

In response to the question, *"Over the past 7 days, on how many days did you engage in physical activity that increased your heart rate or made you breathe harder for a total of at least 60 minutes per day?"* 7.1% of students from general schools reported engaging in no physical activity. Additionally, 18.2% of general school students reported engaging in physical activity for 5 or more days.

In contrast, 24.7% of students from international alternative schools reported no physical activity in the past week—a significantly higher proportion compared to general schools. However, the proportion of students who engaged in physical activity for 5 or more days was 23.3%, slightly higher than that of general schools.

This indicates that while a greater proportion of students in international alternative schools are entirely inactive, those who are active tend to engage in physical activity more frequently than their counterparts in general schools.

Days of Physical Activity (≥60 mins)	General School (n=52,880)	International Alternative School (n=73)
0 days	3,752 (7.1%)	18 (24.7%)
1-2 days	23,913 (45.2%)	18 (24.6%)
3-4 days	15,556 (29.5%)	20 (27.4%)
5+ days	9,659 (18.2%)	17 (23.3%)

table. 5 : Physical Activities

<Adequacy of Sleep>

In response to the question, *"Over the past 7 days, do you feel that the amount of sleep you had was sufficient for recovering from fatigue?"* 27.8% of students from general schools answered "not sufficient," the highest proportion among responses, while only 8.0% reported "Very sufficient."

Among students from international alternative schools, the proportion who answered "not sufficient" was similar at 28.8%. However, 26.0% of students from international alternative schools reported "sufficient," a notably higher percentage compared to students in general schools.

This suggests that while insufficient sleep is a common issue across both school types, students in international alternative schools are more likely to perceive their sleep as sufficient, reflecting potential differences in lifestyle or time management practices between the two groups.

Sleep Sufficiency Level	General School (n=52,880)	International Alternative School (n=73)
Very Sufficient	4,209 (8.0%)	6 (8.2%)
Sufficient	9,802 (18.5%)	19 (26.0%)
Neutral	17,793 (33.6%)	18 (24.7%)
Not Sufficient	14,696 (27.8%)	21 (28.8%)
Not Sufficient at All	6,380 (12.1%)	9 (12.3%)

table. 6 : Sleep Sufficiency

Discussion & Conclusion

<Summary of the Study>

This study analyzed the correlations between stress, physical activity, and sleep among students from general schools and international alternative schools in South Korea. By comparing the differences between the two groups, this study aimed to derive practical implications for reducing adolescent stress. For this purpose, data across the Youth Health General schools utilized behavior surveys, while six international alternative schools in Yongin, Jeju, Seoul, and other cities conducted surveys based on the same questionnaire items. Items related to stress, physical activity, and sleep were selected for analysis. The collected data were analyzed using IBM SPSS Statistics 22, comparing the results of general schools and international alternative schools.

The findings revealed significant differences between the two groups. Students from international alternative schools identified "academic performance and career concerns" (53.4%) as the primary source of stress, a much higher rate than their peers in general schools (24.7%). Additionally, stress related to appearance was more prevalent among students in international alternative schools (15.1%) compared to general schools (4.0%). Regarding stress perception, the majority of general school students reported "feeling a little stressed" (45.1%), whereas a higher percentage of international alternative school students reported "feeling very stressed" (13.7%).

In terms of happiness, international alternative school students showed higher levels of subjective happiness, with 27.4% reporting they were "very happy" and 34.2% reporting they were "somewhat happy," compared to 21.5% and 36.5%, respectively, for general school students. Regarding physical activity, 7.1% of general school students reported no physical activity in the past week, with 18.2% engaging in physical activity for 5 or more days. In contrast, 24.7% of international alternative school students reported no physical activity, while 23.3% reported engaging in physical activity for 5 or more days. For sleep adequacy, both groups had a high proportion of students who felt their sleep was "not sufficient" (27.8% in general schools, 28.8% in international alternative schools). However, 26.0% of international alternative

school students reported their sleep was "sufficient," compared to 18.5% of general school students.

Despite experiencing higher levels of stress and burden, students in international alternative schools reported more positive outcomes in happiness and physical activity, as well as relatively higher satisfaction with sleep.

<Discussion>

The results indicated that students in international alternative schools reported higher subjective happiness. Among these students, 27.4% responded they were "very happy," compared to 21.5% in general schools. This suggests that the flexible learning environment, individualized support, and diverse, enjoyable activities in international alternative schools contribute to greater emotional satisfaction. However, international alternative school students also reported higher levels of perceived stress, with a higher proportion (13.7%) feeling "very stressed." This may stem from the high academic expectations and rigorous curricula, such as AP and IB programs, which likely increase psychological pressure. Thus, additional support strategies are necessary to balance academic demands with happiness in international alternative schools.

The main stressors differed significantly by school type. General school students identified "peer and senior/junior relationships" (34.2%) and "academic performance and career concerns" (24.7%) as their primary sources of stress, reflecting the combined pressures of interpersonal relationships and competitive academic environments. Conversely, international alternative school students overwhelmingly cited "academic performance and career concerns" (53.4%) as their main stressor, likely due to the high expectations and international standard curricula. Interestingly, stress related to appearance was also more prevalent in international alternative schools (15.1%), possibly due to the influence of diverse cultural backgrounds and comparisons among students.

Differences in physical activity levels were observed between the two groups. While 23.3% of international alternative school students engaged in physical activity for 5 or more days, compared to 18.2% in general schools, a higher proportion of international alternative school students (24.7%) reported no physical activity at all compared to general school students (7.1%). This suggests that while international alternative schools promote voluntary participation in activities such as sports clubs, the lack of structured, mandatory programs may delay engagement. These findings highlight the need for schools to develop inclusive programs that ensure equal opportunities for all students to engage in physical activity. Additionally, students who engaged in physical activity for 5 or more days reported lower stress levels, reaffirming the role of physical activity in alleviating adolescent stress.

Sleep satisfaction varied depending on the school environment and academic operations. A higher proportion of international alternative school students (26.0%) reported their sleep as "sufficient" compared to general school students (18.5%). This may reflect the flexible schedules and student-centered environments in international alternative schools, which contribute to

improved sleep quality. However, 41.1% of international alternative school students still reported insufficient sleep, indicating room for further improvements in providing an environment conducive to adequate rest. In general schools, higher rates of sleep insufficiency were likely due to excessive academic workloads and long study hours. Schools should consider adjusting academic schedules and curricula to allow students more time for rest, as insufficient sleep negatively impacts stress levels and academic efficiency.

<Recommendations>

Based on the findings, several recommendations can be made to address the identified issues and improve adolescent well-being. First, educational environments should be improved, particularly in general schools, by adopting the flexible approaches seen in international alternative schools. This includes creating environments where students can prioritize holistic development rather than solely focusing on academic performance. Reducing the emphasis on grades in evaluations and encouraging experiential and creative learning activities would foster a more balanced and enriching educational experience.

Second, tailored stress management programs should be implemented to address the specific stressors experienced by students in different school types. For international alternative schools, enhancing counseling services to help alleviate academic pressure could be highly effective. In general schools, introducing activities designed to prevent school violence and improving peer relationships would be beneficial. These targeted interventions would ensure that stress management strategies are relevant and impactful.

Finally, schools should actively promote physical activity and sufficient sleep as essential components of stress reduction. Regular exercise programs and sleep education should be implemented to help students understand the importance of these factors for their physical and mental health. Such initiatives would not only alleviate stress but also support adolescents in achieving balanced growth and maintaining overall health. These efforts are critical for fostering a supportive school environment that prioritizes the well-being of all students.

<Limitations and Future Research>

While this study highlighted differences between school types and confirmed the importance of physical activity and sleep in reducing stress, it should note several limitations. The significant disparity in sample sizes between general and international alternative schools may limit the generalizability of the findings. Additionally, the reliance on self-reported survey data introduces the potential for subjective bias. Future research should include larger and more diverse samples across various school types and employ qualitative interviews for deeper insights.

Works Cited

- Statistics Korea. Status of Housing and Land Ownership. Statistics Korea, Accessed 2 Nov. 2024, kostat.go.kr/boardDownload.es?bid=12305&list_no=428617&seq=2.
- Ministry of Health and Welfare. Measures to Strengthen Health Insurance Coverage. Ministry of Health and Welfare, Accessed 2 Nov. 2024, www.mohw.go.kr/board.es?mid=a10503000000&bid=0027&list_no=1483270&act=view&.
- Kim, Jisoo. "The Impact of Corporate Social Responsibility on Organizational Commitment." *Korean Journal of Business Ethics*, vol. 10, no. 2, 2014, pp. 45–67. Accessed 2 Nov. 2024, www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART001869904
- Statistics Korea. Social Survey Report: The Life of Koreans. Statistics Korea, Accessed 2 Nov. 2024, kostat.go.kr/boardDownload.es?bid=12306&list_no=370068&seq=1.
- Seo, Sung-Ik, and Soo-Yeon Noh. "Changes in Physical Activity Frequency of Korean Adolescents after COVID-19 and Its Effects on Stress, Sleep Fatigue Recovery, Loneliness, and Generalized Anxiety Disorder." *The Humanities and Social Sciences* 21, vol. 13, no. 6, 2022, pp. 1553–1568.
- Yoon, Jung-Min, Hye-Ran Lee, and Geon-Hee Lee. "A Clinical Review of Sleep Disorders in Children and Adolescents." *Annals of Child Neurology*, vol. 15, 2007, pp. 185–191.
- "Factors Influencing the Learning Attitudes and School Satisfaction of Students in Global Alternative Schools." Master's Thesis, Graduate School of Education, Yonsei University, 2023, pp. 5–8.
- Jung, Yoo-Sung. *What Is Alternative Education?* 1997.
- Seo, Dam-I. "Factors Influencing Learning Attitudes and School Satisfaction of Students in International Alternative Schools." Master's Thesis, Graduate School of Education, Yonsei University, 2023, pp. 5–8.
- Nam, Jin-Seok. "A Qualitative Study on the Culture of International Alternative Schools." Master's Thesis, Graduate School of Public Administration, Korea University, 2009.
- Kang, Dae-Yong. "A Qualitative Study on the Educational Experiences of Graduates from International Alternative Schools." Master's Thesis, Seoul National University, 2018.
- Seo, Dam-I. "A Qualitative Case Study on the School Education Experiences of Students in International Alternative Schools." Master's Thesis, Graduate School, Korea University, 2017, Seoul.
- Ham, Young-Joo. "A Qualitative Study on the Educational Outcomes of Graduates from Christian Alternative Schools: Focusing on Dilemma Narrative Analysis." *Theology and Praxis*, vol. 47, 2015, pp. 459–483.
- Kim, Junghyun, Sungbyuk Kim, and Inkyung Jung. "The Effects of Academic Stress on School Life Adjustment in Adolescents." *Korean Journal of Youth Studies*, vol. 25, no. 4, 2014, pp. 241–269.

- Kim, Junghyun, Sungbyuk Kim, and Inkyung Jung. "The Effect of Academic Stress on School Adjustment in Adolescents: Focusing on the Mediating Effect of Stress Coping." *Studies on Korean Youth*, vol. 25, no. 4, 2014, pp. 241–269. National Youth Policy Institute Repository, www.nypi.re.kr/repository/handle/2022.oak/2032.
- Jung, Mi-Ra, and Eun Jeong. "Factors Influencing Sleep among Korean Adolescents: Using Data from the 7th-8th Korea National Health and Nutrition Examination Survey, 2016-2020." *Journal of Korean Biological Nursing Science*, vol. 24, no. 3, 31 Aug. 2022, pp. 181–189, <https://doi.org/10.7586/jkbns.2022.24.3.181>.
- Kim, Yeon-Sook, and Eun Kim. "The Study on the Moderating Effect of Breakfast, Sleep, and Physical Activity on the Relationship between Stress and Quality of Life in Children and Adolescents." *Journal of the Korean Society for Wellness*, vol. 16, no. 4, 30 Nov. 2021, pp. 201–212, <https://doi.org/10.21097/ksw.2021.11.16.4.201>.
- Seo, Kyung-Hee. "A Study on Stress, Depression, and Suicidal Ideation in High School Male Students: A Comparison by School Type." Master's Thesis, Inha University, 2016, pp. 10–13.
- Kim, Junghyun, Sungbyuk Kim, and Inkyung Jung. "The Effect of Academic Stress on School Life Adjustment in Adolescents: Focusing on the Mediating Effect of Stress Coping." *Korean Journal of Youth Studies*, vol. 25, no. 4, 2014, pp. 241–269.
- Chakrabarti, Reeta. "South Korea's Schools: Long Days, High Results." *BBC News*, 2 Dec. 2013, www.bbc.com/news/world-asia-25239189.
- Jarvis, Jonathan A., Allison W. Corbett, Jared D. Thorpe, and Mikaela J. Dufur. "Too Much of a Good Thing: Social Capital and Academic Stress in South Korea." *Social Sciences*, vol. 9, no. 11, 2020, p. 187, <https://doi.org/10.3390/socsci9110187>.
- "South Korea's Industrial Policy: Growth with Inefficiency." National Bureau of Economic Research, 1 Nov. 2021, www.nber.org/digest/202111/south-koreas-industrial-policy-growth-inefficiency.
- OECD. "Suicide Rates." OECD Data, data.oecd.org/healthstat/suiciderates.html

Repairing the Cracked Foundation of Financial Equality: Policy Solutions to Reverse the Legacy of Housing Discrimination By Anthony Borrelli

It was the early 1960s, and African American veteran Clyde Ross was looking to move his young family into a decent Chicago neighborhood. He was financially qualified to do so. Yet, when Ross attempted to secure a mortgage, the bank informed him that no such option was available (Coates). It was because of his race, and Ross was not the only one. In fact, from the 1930s through most of the 1960s, federal agencies denied government-backed home mortgages to African Americans across the United States due to the color of their skin. This practice, commonly referred to as *redlining*, was one of many discriminatory housing practices that severely restricted the homebuying efforts of Black Americans during the 20th century, confining them to separate neighborhoods from their white counterparts. Similarly, many cities implemented racially motivated zoning laws to prevent the integration of Black residents into white communities, designating less favorable localities for Black Americans. Facing pressure from the growing civil rights movement, President Lyndon B. Johnson enacted the Fair Housing Act in 1968, which officially outlawed all discriminatory housing practices, including redlining and racial zoning policies. Yet, inadvertently, more hidden forms of racial housing discrimination gained prominence, including *racial steering*—the directing of homebuyers to racially similar neighborhoods by real estate agents—and the undervaluing of Black homes and communities by a majority-white appraisal industry. Today, as roughly 49% of Black buyers face unequal treatment by the real estate industry, housing discrimination remains a pressing inequality in modern American society, not simply a wrong of the past (Choi et al.).

Without access to a mortgage, Ross resorted to buying a house on contract, a type of payment plan that severely disadvantages the buyer. Consequently, he had to take on three jobs to finance his home and support his children (Coates). For many Black Americans who were zoned into inadequate neighborhoods or refused mortgages because of their race, their financial situations suffered as well. Zoning laws that concentrated Black Americans in substandard areas and redlining that blocked home loans to African Americans prevented many from acquiring and passing down equal home equity as whites, limiting the wealth of future generations. Moreover, these practices led to a lasting lack of demand in communities with large Black populations, causing African American wealth from homeownership to remain significantly low—a disparity that is further exacerbated by modern racial steering and biased appraisals (Zonta). Largely as a result, the median white family income is roughly 60% greater than that of Black families today (Rothstein 184). Thus, historic housing policies such as racial zoning and redlining, combined with the current existence of racial steering and appraisal bias, have disproportionately limited the home wealth of many Black Americans, fueling the current financial gap between Black and white Americans; such financial inequality must be remedied by the national increase of affordable housing in more affluent neighborhoods, the diversification of the appraisal industry, and government reparations for redlined Black individuals and their descendants, as the

responsibility of the United States government is to ensure equal opportunities for all its citizens and to rectify its past injustices.

Following the emancipation of enslaved people in 1865, many communities across the United States began to create laws and policies to keep African Americans separate from whites, including zoning laws that segregated neighborhoods and deliberately harmed Black residents. After gaining their freedom, many formerly enslaved individuals began to move out of the former Confederate South to escape racial violence and find work in the North and West. This movement accelerated in the early 20th century, as the advent of World War I created more industrial opportunities for hundreds of thousands of Black Americans in the North. Consequently, white Americans in these areas, fearing the integration of their communities, began to clamp down on the residential freedom of Black Americans. For example, many city governments began to “adopt zoning rules decreeing separate living areas for Black and white families” to “isolate white families in all-white urban neighborhoods” (Rothstein 44). By voting for such zoning laws, white residents hoped to “save [their] home[s]” from a “negro invasion,” as warned about in a 1915 Missourian leaflet in support of zoning segregation (United Welfare Association). In addition to imposing residential segregation, zoning laws also ensured that the neighborhoods where Black individuals lived were severely disadvantaged. In particular, many cities across the United States in the early 20th century “changed an area’s zoning from residential to industrial if African American families had begun to move into it,” placing unfavorable industries and toxic waste sites in predominantly Black neighborhoods (Rothstein 50). Similarly, many cities allowed for “taverns, liquor stores, nightclubs, and houses of prostitution” in largely Black localities, yet they “prohibited these as zoning violations in neighborhoods where whites lived” (Rothstein 50). Through such racist legislation, majority-white city governments selfishly promoted the deterioration of Black areas for the benefit of white communities, reflecting the systemic treatment of African American residents as inferior to whites. Hence, racial zoning laws legally enforced residential segregation and the decline of Black areas into slums.

Subsequently, beginning during the New Deal, federal redlining practices deliberately denied government-backed home mortgages to neighborhoods with large numbers of African Americans, significantly limiting their homebuying opportunities. During the Great Depression, there was a lack of housing availability across the United States. In response, President Franklin D. Roosevelt established the Home Owners' Loan Corporation (HOLC) and the Federal Housing Administration (FHA) to refinance mortgages and provide federal insurance on home loans, respectively (Zonta). To determine which communities were safe for mortgage services, the HOLC developed color-coded maps, marking the most “hazardous” neighborhoods in red, while coloring the “best” in green (Meier and Mitchell). Yet, this system was heavily based on race; the HOLC often deemed areas with significant Black populations as “hazardous” while labeling majority white areas as “best” as African Americans were prejudicially viewed as financial threats to lenders (Singletary). For instance, in 1939, HOLC officials visited a neighborhood in Cleveland, Ohio, that was 55% Black, assigning it a “hazardous” rating; in the report, the

officials documented “strong colored infiltration” and a “detrimental change of ownership occupancy from white to colored” as their reasons behind this rating, exemplifying the use of racial factors in these decisions (Best and Mejía). After the HOLC created these loan security maps, the FHA used them to determine which communities to provide home loan insurance to, insuring home loans to those in the “best” neighborhoods, while refusing such insurance to those in “hazardous” areas. Since those marked “hazardous” were largely Black in population, the FHA systemically denied government-backed home loans to African Americans across the United States, just like Clyde Ross. Thus, through the racist lending insurance services established during the New Deal, the United States government institutionally prevented opportunities for Black Americans to purchase new houses, keeping them isolated. This practice would remain in place until 1968, when Lyndon B. Johnson enacted the Fair Housing Act, officially outlawing redlining and all other forms of housing discrimination.

Yet, after the passage of the Fair Housing Act, more covert forms of housing discrimination gained prominence, including the practice of racial steering, which continues to reinforce segregation and home inequality between Black and white Americans today. Although the Fair Housing Act deemed such discrimination illegal, it had the effect of “push[ing] a lot of these practices underground,” according to Associate Professor Jacob Faber, who specializes in racial inequality and housing research (Choi et al.). One of the most common of these “underground” methods is steering; according to the National Association of Realtors, 39% of Black homebuyers in 2023 reported being steered toward certain neighborhoods due to their race—a number that increased 12% since 2022 (“Snapshot of Race”). To investigate this practice, *Newsday* conducted a three-year study in Long Island, New York—beginning in 2019—in which 25 undercover testers pretended to look for homes with 93 real estate agents. In one of these tests, a Realty Connect USA agent recommended areas that were on average 88% white to the white tester, while suggesting neighborhoods that were on average 58% white for the Black tester. Both specified that they were looking for houses that were \$600,000 and within 60 minutes of New York Penn Station (Choi et al.). This racial composition difference of 30% exemplifies how real estate agents can provide drastically different services to their clients simply based on racial factors, segregating white and Black homebuyers into racially disparate areas. While racial steering limits both Black and white Americans in their housing options, the areas that Black individuals are steered toward typically have “higher poverty levels...and lower housing quality compared with neighborhoods to where whites relocate,” disproportionately putting Black Americans at a greater disadvantage in their housing choices (Zonta). Therefore, steering continues segregation patterns while unfairly restricting Black Americans’ residential opportunities.

Altogether, both historic and modern housing discrimination practices have prevented many Black Americans from acquiring and passing down home wealth, contributing significantly to the racial wealth gap today. Since the FHA’s racist lending insurance policies barred many African Americans from purchasing new homes, they could not participate in the accumulation of home equity that their white counterparts had access to. As whites began to buy

new homes, their properties increased in value and provided substantial wealth in return, while redlined Black Americans were left out of such opportunities (Gross). Since equity is typically passed down to future generations, many African Americans continue to lack home equity in comparison to whites today (Rothstein 185). Redlining policies also created a persistent lack of investment, employment opportunities, and demand in many Black localities, especially as industrial racial zoning transformed many into undesirable slums (Zonta). Without sufficient equity and property demand, countless Black Americans are restricted to continue buying in Black neighborhoods with a similar lack of financial opportunities, remaining trapped in a vicious cycle of financial disadvantage (Zonta). This cycle, combined with the steering of African Americans into Black areas, continues to trap many Black Americans in financially inferior locations with worse employment and educational opportunities; as a result, the average Black American resides in a neighborhood that is only 35% white and typically 23% lower in value than non-Black communities (Rothwell and Perry; Zonta). These statistics expose the severity of persisting residential segregation that continues to contain Black Americans in less affluent neighborhoods. Likewise, according to data from the real estate company Zillow between 1996 and 2018, the median value of homes in previously “best” areas grew 230.8% to \$640,238, whereas the median value in “hazardous” communities grew 203.1% to only \$276,199 (Singletary). This connection between previously redlined neighborhoods and their significant lack of home appreciation and value today displays how the FHA’s redlining practices still contribute to the financial decline of the Black areas it once restricted. Largely as a result, the average home wealth of white Americans was \$215,800 in 2016, while only \$94,400 for Black homeowners (Zonta). This striking disparity reveals the detrimental financial ramifications of racial zoning, redlining, and steering that current generations of Black Americans continue to suffer from, while many whites conversely prosper off of their homes. Since home equity is a significant component of one’s wealth, the discriminatory housing policies responsible for such disparities are thus notable factors of the racial wealth gap today.

Additionally, due to racial biases, current appraisers often undervalue homes in predominantly Black neighborhoods, along with Black-owned homes in other areas, further worsening the financial standings of African American homeowners. A 2022 study by the Brookings Institution, a nonpartisan public policy research organization, found that appraisers tend to undervalue roughly 10% of homes in Black communities at prices 15% lower than expected. This same study found that undervaluing occurs 1.9 times more in majority-Black neighborhoods than it does in those that are majority-white (Rothwell and Perry). While an undervaluation of 15% may appear minimal, this statistic can translate to a loss of tens of thousands of dollars, robbing Black homeowners of their deserved wealth because of racial prejudice. Further, although these statistics only regard majority-Black neighborhoods, many Black homeowners in white communities similarly have their homes undervalued by appraisers. For example, in 2023, Black couple Erica and Aaron Parker in majority-white Loveland, Ohio, expected to sell their home for \$525,000; yet, the appraiser valued their home at \$60,000 less. The family then hired a new appraiser, this time hiding all family photos and items that would

reveal they were Black, having their white neighbor sit in for the valuation instead. The appraisal then returned at a value roughly \$92,000 greater than the first (Ellis). Since home valuations are meant to be formulaic, this drastic difference in value between the two appraisals was more plausibly caused by racial bias than by any other factor, exposing the severe monetary losses that a racist valuation can inflict on African Americans, while also revealing the disproportionate financial benefit of white owners of similar—if not the same—properties. Given that many Black Americans are already lacking in home equity, racist appraisals work to minimize further the wealth they can generate from their homes, amplifying the racial wealth gap.

To remedy existing residential segregation that confines many Black Americans to substandard communities, the federal government must mandate systems of affordable housing in wealthier locations for all states. Currently, most affordable housing units across the United States are placed in lower-income neighborhoods due to cheaper land costs. As a result, such housing “concentrates poverty and continues the cycle of racial and socioeconomic segregation set in motion by racialized zoning practices,” redlining, and racial steering (Fatheree). Therefore, for many Black Americans in need of affordable housing, the current system keeps them trapped in the same financially inferior areas that discriminatory housing practices have already restricted them to. However, placing affordable housing projects in affluent communities would instead encourage movement to locations with greater economic opportunities, allowing these same Black residents to improve their financial situations and escape the grasp of persisting residential segregation. A current example of a state law that promotes the creation of affordable housing in wealthier communities is Massachusetts General Laws Chapter 40B; under Chapter 40B, real estate developers are allowed to “override local zoning bylaws in order to increase the number of affordable homes in municipalities where less than 10% of the housing is defined as affordable” (“40B Affordable”). By overriding zoning laws in towns and cities that are largely unaffordable, developers can build affordable housing units where they otherwise would not have been located, such as in wealthier, suburban areas. Indeed, according to the Turner Center for Housing Innovation at UC Berkeley, Chapter 40B “has been successful in increasing the number of units affordable to low- and moderate-income housing built across Massachusetts, including in affluent suburbs” where such housing is typically not found (Reid et al.). Given the success of Chapter 40B in Massachusetts, the implementation of a similar program in every state would increase the number of affordable homes in affluent communities on a national level, allowing for the relocation of financially disadvantaged Black residents across the country. For such a program to exist throughout the nation, the federal government must require affordable housing systems like Chapter 40B in all states, fulfilling the American ideal that every citizen has access to equal financial opportunities.

Yet, many residents in affluent areas are opposed to affordable housing in their communities out of fear that their property values will suffer. This concern is reasonable for homeowners, given the importance of home wealth and equity on their total financial well-being. However, the Turner Center for Housing Innovation at UC Berkeley has found that the affordable housing projects created under Massachusetts General Laws Chapter 40B have had “little to no

impact on the surrounding community” in terms of home values (Reid et al.). Thus, if the federal government were to require a system of affordable housing in all states modeled off of Chapter 40B, housing values in affluent communities would likely remain unaffected as well. Without a threat to home prices, the construction of affordable housing in wealthier communities across the United States provides a promising solution to increase residential opportunities for Black victims of residential segregation.

Additionally, the appraisal industry must diversify its workforce to reduce racially biased appraisals that further disadvantage Black homeowners. Currently, the industry is almost completely racially homogenous, with a makeup that was 92% white and 4% Black as of 2022 (Ellis). Given the regularity of racially biased appraisals previously stated—for about 10% of homes in Black neighborhoods—this lack of diversity plausibly contributes to the prominence of racial prejudice in the field. For this reason, the industry needs a more diverse set of appraisers, as an increase in Black staff would likely reduce the possibility of Black homes being undervalued due to racial bias. One method for appraisal businesses to achieve this diversity is actively recruiting Black individuals. Currently, the National Association of Real Estate Brokers holds annual events at historically Black colleges and universities to promote the appraisal industry as a career option for Black students to consider (Ellis). Following this model, appraisal businesses throughout the United States should host events in areas with large numbers of Black students and residents to raise their awareness of the industry and encourage these individuals to join. Additionally, in 2021, President Joe Biden’s administration created the Action Plan to Advance Property Appraisal and Valuation Equity (PAVE), a task force that has since been investigating the appraisal industry and providing recommendations on how it can diversify its workforce. For example, PAVE recommends that the industry lower its entry requirements to increase Black participation, such as experience requirements and the need for a college degree, finding that there “is little existing support that a college degree makes appraisers more ethical, accurate, or credible” (“Action Plan”). Since these requirements unreasonably discourage many African Americans from joining the field, as college degrees are not crucial for appraisers to provide quality service, appraisal businesses must lower these barriers to increase employee diversity. Not only would this action decrease biases that lead to the undervaluation of Black homes, but it would also build a credible reputation with consumers that would reciprocally benefit the company. Thus, the diversification of the appraisal industry offers an overall advantageous solution to counteract racially prejudiced property valuations.

Likewise, the federal government must mandate racial bias training for all appraisers across the United States. As of now, only twelve states, including New York, Illinois, and California, require racial training for appraisers (Wheaton). With racial bias still devaluing many Black homes across the United States, the lack of state-required racial training in the majority of the country is likely a significant reason that prejudice continues to skew many appraisals. Hence, a major step toward eliminating racial bias in the appraisal industry would be for the federal government to require such training in all fifty states so that every appraiser could be equally informed on how to provide fair treatment to all homeowners. Since these training

requirements already exist in twelve states, the federal government could easily model its mandate off these current laws. However, some appraisal business owners and employees may feel frustrated about spending the necessary time and resources to undergo mandatory training; such feelings are rational, as businesses could instead spend this time working. Yet, when the goal of the appraisal industry is to provide *completely accurate* property assessments to homeowners, racial training is a necessary step to ensure that these businesses are indeed serving their stated function of providing correct valuations to their clients. Additionally, similar to the effects of a diversified appraisal industry, racial training would also increase consumer trust in these companies, as more clients would feel confident in the accuracy of the appraisals they are given. As appraisers continue to undervalue many Black homes, federally mandated racial training would ensure fair valuations for all homeowners while mutually benefiting the industry itself.

Before delving into the justifications for a federal reparations program, the discussion of compensatory payments for housing discrimination must only focus on redlined Black individuals and their descendants for such a program to best obtain political success. In recent years, the discussion about reparations for African Americans has gained national attention, although it tends to focus on payments for other injuries and inequalities as well, not solely housing discrimination (Kaplan and Valls). For example, California's Task Force to Study and Develop Reparation Proposals for African Americans recently devised a plan to compensate the state's Black population for the unjust financial effects of slavery, mass incarceration, health disparities, Jim Crow laws, and redlining; this plan would cost the state roughly \$800 billion if implemented (Concha). While payments for all of these injuries are still important to consider, the multitude of factors listed and the large cost of the entire package make any possibility of reparations too extensive for many to digest, becoming vulnerable to political opposition. Many political critics also find the issue of slavery too far in the past for the federal government to readdress today (Kaplan and Valls). Therefore, within the discussion of residential discrimination, the case for Black reparations must solely focus on compensating for the damages caused by racist housing policies, which are both recent and specific. By honing in the task further to only include reparations for Black victims of redlining and their descendants, the argument for compensation becomes even more politically palatable and financially manageable, especially given that many of the HOLC redlining maps still exist and could be used to identify past redlined residents, their descendants, and an estimate of their financial losses. With this specific purpose, many voters would feel much more comfortable putting their support behind a reparations program for African Americans.

Now, with the focus on redlined Black individuals and their descendants, the United States government must provide reparations to redress the financial inequalities that its racist lending policies created. As previously displayed, the government's redlining policies from the 1930s through most of the 1960s significantly limited the abilities of many Black Americans to build up and pass down equal home equity to their children compared to whites, heavily driving the racial wealth gap today. Through a reparations program, the federal government would

greatly reduce these wealth disparities that its policies are responsible for, promoting greater financial equality between both races. Indeed, a model for the United States to follow already exists; in the 1950s, West Germany agreed to provide reparations to Israel for the atrocities it committed toward the Jewish people during the Holocaust, which totals more than \$7 billion today (Coates). Germany's payments to the Jewish people demonstrate to the United States that a reparations program is indeed achievable, with the framework already available. While reparations could never make up for the losses incurred through mass murder, they did result in the tripling of Israel's GNP and the creation of 45,000 new jobs (Coates). Given the substantial economic growth that these payments produced, a reparations program for redlining in the United States would likely have similar success in bettering the financial standing of many Black Americans, serving as a promising solution to lowering the racial wealth disparities that government-enforced housing discrimination created. Reparations could also stimulate the economy of the United States as it did for Israel, as more Black Americans would have greater spending power and financial mobility. If Germany can take proactive steps to reverse the damage of a devastating genocide, then the United States must be able to remediate wealth disparities that it is directly responsible for as well. For a true democracy to function, a government must be able to take accountability for its injustices and rectify them, actively displaying its commitment to protect the equality and opportunity of all its citizens. Without reparations payments, the federal government continues to allow countless Black individuals to suffer at the hands of its past policies; yet, by following in the footsteps of a sympathetic and responsible Germany, the United States government would properly fill the cracks it created in the foundation of Black wealth.

However, a largely criticized limitation of federal reparations to Black Americans is that they could set a precedent for all other minority groups to claim compensation for their own systemic injuries. Since its beginning, the United States has endured a complicated history steeped in racism toward almost all non-white groups, not just Black individuals. For this reason, if the federal government provided redlining reparations for African Americans, many other racial minorities may feel deserving of reparations as well, considering that they, too, have been oppressed by the United States government. Regarding this challenge that reparations for Black Americans presents, Republican political commentator David Frum writes,

With any program of reparations...other claimants will come forward....[W]hat about Native Americans, who lost a whole continent? What about Mexican-Americans, who were deprived by the Mexican-American war of the right to migrate into half their former country? Japanese Americans, interned during World War II? Chinese Americans, the victims of coolie labor and the Oriental Exclusion Acts? (Frum)

While his conservative political beliefs may influence his opinion against reparations in favor of a smaller, less-involved government, Frum's concern about what reparations for Black

Americans could mean for other minority groups is reasonable, considering that payments for all injured parties would become a complex and expensive task. However, by introducing other racial groups, Frum and similar opponents of Black reparations distract from the original discussion, which is solely focused on remediating the highly specific financial damages of redlining policies for African Americans. Concerns about how other minority groups may respond to their own injuries are irrelevant to the particular conversation regarding reparations for redlining, as they only take away from addressing the plight of Black Americans. Thus, the raising of these distractions is largely an oppositional tactic to deflect attention from the possibility of compensation. Furthermore, by repeatedly presenting hypothetical cases of groups that may demand reparations in a series of rhetorical questions, Frum overcomplicates the argument for Black reparations, fatiguing its potential proponents and making reparations for Black redlining victims appear impossible. Lastly, this argument is grounded in the assumption that compensation for other minorities is objectively an issue that needs to be prevented. Yet, these groups may very well deserve their own reparations, as long as the conversations regarding those reparations occur independently, without being used to block potential remedies for African Americans. Therefore, concerns about this potential impact on other minority communities must not interfere with the discussion about redlining reparations for Black Americans.

Ultimately, the legacy of discriminatory housing policies continues to contribute significantly to the wealth gap between Black and white Americans, justifying the need for affordable housing opportunities in opportunity-rich communities, the diversifying and increased racial education of appraisers, and federal compensation for African American redlining victims and their descendants. Although outlawed today, citywide racial zoning policies combined with federally enforced lending discrimination worked to limit the home wealth of Black Americans for generations, while contrarily allowing white Americans to benefit from a lasting appreciation in home values and accumulation of equity. Today, these racial disparities in home wealth are continually worsened by the steering of African Americans into poorer Black neighborhoods and the pervasiveness of bias in the appraisal industry. Since a significant portion of the wealth gap between Black and white Americans is attributable to these existing housing inequities, Black Americans must finally have access to equal and fair residential opportunities throughout the United States. A nationwide requirement of affordable housing programs directed at wealthier communities would allow countless Black Americans to relocate out of government-made slums, while the diversification and racial training of appraisers would help reduce persisting racial bias in the housing industry. With the government's redlining practices largely responsible for the current financial struggles of Black individuals today, reparations would responsibly restore the economic opportunities that the federal government has robbed many African Americans of for generations.

Clyde Ross passed away two years ago. While he lived to witness significant progress toward housing equality, such as the Fair Housing Act and President Biden's PAVE, Black Americans continue to face a racially biased housing market and its devastating financial

consequences. With Black Americans still suffering the financial injuries of residential inequality that Clyde Ross once faced, both the federal government and private industry must take proactive steps to reverse the harmful legacies of their own racist housing policies. In a democracy where all citizens are promised equal economic opportunities, the federal government must be held accountable for the financial inequities it has created amongst its population, returning the wealth and residential freedom that its redlining policies have stripped from Black Americans. Concerning private forms of housing discrimination, appraisal businesses must ensure the unbiased treatment of their clients to promote racial equity in the real estate market, further working to close the racial wealth gap. Yet, when the private sector fails to do so, the government must serve as the ultimate legal entity to enforce racial fairness, fostering the equal and just society it has promised its citizens since its creation. Once countless Black Americans obtain fair access to residential and financial opportunities, they will finally be able to open the door to the American Dream that has been sealed shut for generations.

Works Cited

- "40B Affordable Housing for Homebuyers." *NeighborWorks Housing Solutions*, nhs.mass.gov/chapter-40b-affordable-housing/. Accessed 21 May 2024.
- "Action Plan to Advance Property Appraisal and Valuation Equity." PAVE, U.S. Department of Housing and Urban Development, pave.hud.gov/actionplan. Accessed 30 May 2024.
- Best, Ryan, and Elena Mejía. "The Lasting Legacy of Redlining." *FiveThirtyEight*, ABC, 9 Feb. 2022, projects.fivethirtyeight.com/redlining/#explore-top. Accessed 20 May 2024.
- Choi, Ann, et al. "Long Island Divided." Edited by Arthur Browne. *Newsday*, 17 Nov. 2019, projects.newsday.com/long-island/real-estate-agents-investigation/#open-paywall-message. Accessed 21 May 2024.
- Coates, Ta-Nehisi. "The Case for Reparations." *The Atlantic*, June 2014, www.theatlantic.com/magazine/archive/2014/06/the-case-for-reparations/361631/. Accessed 29 Apr. 2024.
- Concha, Joe. "Reality Finally Comes for California's Reparations Plan." *The Hill*, 14 May 2023, thehill.com/opinion/civil-rights/4003151-reality-finally-comes-for-californias-reparations-plan/. Accessed 31 May 2024.
- Ellis, Nicquel Terry. "A Black Family Says They 'Whitewashed' Their Home to Get a Higher Appraisal. They're Not the Only Ones." *CNN*, 20 Apr. 2023, www.cnn.com/2023/04/15/us/real-estate-appraisals-bias-reaj/index.html. Accessed 29 Apr. 2024.
- Fatheree, George. "A Brief History of Racial Zoning and How to Reverse the Lasting Effects of Housing Discrimination." *Urban Land*, 20 Feb. 2024, urbanland.uli.org/a-brief-history-of-racial-zoning-neighborhood-associations-and-municipal-zoning. Accessed 21 May 2024.
- Frum, David. "The Impossibility of Reparations." *The Atlantic*, 3 June 2014, www.theatlantic.com/business/archive/2014/06/the-impossibility-of-reparations/372041/. Accessed 29 Apr. 2024.
- Gross, Terry, host. "A 'Forgotten History' of How the U.S. Government Segregated America." *Fresh Air*, 3 May 2017. NPR, www.npr.org/2017/05/03/526655831/a-forgotten-history-of-how-the-u-s-government-segregated-america. Accessed 28 May 2024.
- Kaplan, Jonathan, and Andrew Valls. "Housing Discrimination as a Basis for Black Reparations." *Public Affairs Quarterly*, vol. 21, no. 3, July 2007, pp. 255-73, www.jstor.org/stable/40441462?read-now=1#page_scan_tab_contents. Accessed 29 Apr. 2024.
- Meier, Helen C.S., and Bruce C. Mitchell. "Tracing the Legacy of Redlining: A New Method for Tracking the Origins of Housing Segregation." *National Community Reinvestment Coalition*, Feb. 2022, ncrc.org/redlining-score/. Accessed 20 May 2024.
- Reid, Carolina K., et al. "Borrowing Innovation, Achieving Affordability: What We Can Learn from Massachusetts Chapter 40B." *Turner Center for Housing Innovation*, U of California

- Berkeley, Aug. 2016,
 ternercenter.berkeley.edu/wp-content/uploads/pdfs/California_40B_Working_Paper.pdf.
 Accessed 30 May 2024.
- Rothstein, Richard. *The Color of Law: A Forgotten History of How Our Government Segregated America*. A Liveright paperback ed., New York City, Liveright, 2018.
- Rothwell, Jonathan, and Andre M. Perry. "How Racial Bias in Appraisals Affects the Devaluation of Homes in Majority-Black Neighborhoods." *Brookings*, 5 Dec. 2022, www.brookings.edu/articles/how-racial-bias-in-appraisals-affects-the-devaluation-of-homes-in-majority-black-neighborhoods/. Accessed 29 Apr. 2024.
- Singletary, Michelle. "Being Black Lowers the Value of My Home: The Legacy of Redlining." *The Washington Post*, 23 Oct. 2020, www.washingtonpost.com/business/2020/10/23/redlining-black-wealth/. Accessed 29 Apr. 2024.
- "Snapshot of Race and Home Buying in America." *National Association of Realtors*, www.nar.realtor/research-and-statistics/research-reports/snapshot-of-race-and-home-buying-in-america#:~:text=Experience%20of%20Racial%20Discrimination%20in,less%20than%201%25%20last%20year. Accessed 21 May 2024.
- United Welfare Association. *Look at These Homes Now!* 1915. *Missouri Historical Society*, mohistory.org/collections/item/D02323. Accessed 20 May 2024.
- Wheaton, David. "Fighting Appraisal Bias: How the Government and Housing Industry Can Better Address This Discriminatory Practice." *Legal Defense Fund*, www.naacpldf.org/appraisal-algorithmic-bias-racial-discrimination/. Accessed 21 May 2024.
- Zonta, Michela. "Racial Disparities in Home Appreciation." *Center for American Progress*, 15 July 2019, www.americanprogress.org/article/racial-disparities-home-appreciation/. Accessed 29 Apr. 2024.

Pros and Cons of Digital Media Use as a Coping Strategy By Arjun Chopra

Abstract

With the increased use of digital media over time, people have started to use digital media to help with stress management. This paper will examine the pros and cons of digital media being used as a coping strategy. Research shows social media has the power to relieve people from acute stressors. Further, instant messaging, specifically, was found to improve social connectedness and mitigate stress in some cases. It has also been that digital apps for stress management and mental well-being were effective in improving people's stress and depression levels. Yet, negative effects have been researched as well, which will also be addressed in this paper. Specifically, it has been shown that coping by using digital media can create distractions from aspects of life. Socially, using digital media as a coping strategy can interrupt in-person social interactions, and in turn harm relationships and friendship satisfaction. Digital media use can also end up becoming a distraction from work and academics. This review paper takes a nuanced look at the impacts of using digital media as a coping tool by considering the circumstances in which it can be beneficial while also recognizing its potential harms. Further research with more objective behavior tracking is needed to fully understand the prevalence of this coping strategy, and characterize typical instances of employing it, as well as the short-term and long-term effects.

Introduction

To manage stress, people often employ various coping strategies (i.e., "the thoughts and behaviors used to manage the internal and external demands of a stressful situation" [Folkman]). While, psychologists have categorized these strategies in different ways, much of the literature agrees on distinguishing between beneficial coping strategies (associated with positive outcomes) and maladaptive ones (leading to further problems in the long term) (Carver et al.). Two strategies often considered less useful are behavioral disengagement (i.e., when someone tries to avoid or reduce dealing with the stressor) and denial (Carver et al.). Interestingly, more recent work has begun to acknowledge coping through various forms of digital media with contradicting evidence regarding its effectiveness and potential negative consequences.

People spend more time on digital media than ever before. In 2014, adults in the United States (U.S.) spent an average of 309 minutes per day on digital media, and just ten years later in 2024, adults are spending an average of 508 minutes on digital media, a 64% increase in that time (Guttmann). By 2021, teens had an average screen time of 8 hours and 39 minutes (Moyer). Using digital media can be very distracting and a way to escape external situations, environments, and negative thoughts. This may be because screen-time behavior is associated with an acute decrease in self-reported interoception compared to non-screen-time behaviors (Wallman-Jones et al.) This may suggest that people end up using their devices to disassociate from their surroundings and push away bodily feelings, awareness, and associated emotions (Wallman-Jones et al.).

This paper will examine the pros and cons of coping through digital media use. We will explore the potential benefits, including how digital media use can buffer against acute stress and negative emotions, and provide relief from stress induced by social situations, as well as the role of digital apps designed to decrease both acute and chronic stress. We will then consider the potential drawbacks of using digital media as a coping mechanism because it could become a distraction from various aspects of life including in-person interactions, academics, and the workplace. By examining both sides, we aim to provide a balanced view of digital media use as a coping strategy.

Digital Media's Relief Of Acute and Negative Emotions

Digital media can serve as an effective coping mechanism to alleviate negative emotions and mitigate stress. There is some emerging research investigating how using certain popular social media apps immediately before experiencing an acute stress may be able to diminish both physiological and subjective stress responses to said stressor (Harvie et al; Johnshoy et al; Rus and Tiemensma). For example, one study investigating the impact of social media on acute stress reduction revealed that using Facebook immediately before experiencing an acute social stressor induced by a commonly used experimental paradigm called the Trier Social Stress Test (TSST) can diminish both physiological and subjective stress responses (Rus and Tiemensma). In this study, participants were randomly assigned to two different conditions. The group of participants in the first condition looked at electronic magazines while those in the second condition used Facebook. Following this, the participants in both groups faced an acute social stressor. All participants showed significant increases in subjective and physiological stress markers in response to the stressor. However, when controlling for gender and emotional investment in the website, participants who used Facebook experienced lower levels of psychosocial stress and physiological stress and rated the stressor as less threatening compared to those who did not use Facebook directly before the stressor. There are multiple possible explanations for this observed effect. Unfortunately, a common limitation of these previously conducted studies is that the content of social media consumption was not controlled for. However, future work could consider testing different types of social media content to see which type of content has positive effects and any potential significant differences among types of content. This could help us better understand the mechanisms driving this acute stress reduction effect.

One explanation may be that social media use can increase feelings of social connectedness, support, and general positive social interactions. Numerous pieces of literature discuss how using social media can have positive effects when it is used for social connectedness by building friendships and lowering loneliness (Clark et al; Ryan et al; Winstone et al.). In particular, one aspect of social media, instant messaging, a widespread form of online communication, has been studied for its potential stress-relieving effects. A 2013 study investigated how instant messaging affected adolescents experiencing emotional distress, using a combination of self-report questionnaires, textual analysis, and expert evaluations of conversations (Dolev-Cohen and Barak). The researchers also categorized participants as

introverts or extroverts to obtain more nuanced results. Overall, the study found that instant messaging had a positive impact on the well-being of distressed adolescents. Notably, introverts derived greater benefits from instant messaging compared to extroverts. This finding underscores the variability in digital media's effectiveness as a coping tool, suggesting that its impact may differ based on individual personality traits and specific circumstances.

Other work has continued to highlight individual differences in social media's influence on coping effectiveness. For example, a 2015 study explored how different uses of tablets influence individuals' perceptions of stress reduction (Leung). The research revealed that people with a high desire for solitude and those prone to leisurely boredom were more inclined to believe that solitary tablet use would alleviate their stress. Additionally, individuals who primarily used tablets for social and entertainment purposes anticipated stress-reduction benefits from tablet use. These findings provide preliminary evidence that tablet use can offer stress relief, but importantly, this effect appears to be moderated by certain individual characteristics. The study draws parallels between introverts and those who prefer solitude, noting that both groups reported greater stress reduction from digital media use. This aligns with Dolev-Cohen and Barak's study, where introverts experienced more significant stress relief. Collectively, these studies suggest that certain personality types may derive greater benefits from using digital media as a coping mechanism.

Relief From Averse Social Experiences

There is a lot of research looking at how the negative emotions of social distress, but there is also recent research that highlights certain positive aspects. Digital media can be used to relieve social stress such as anxiety or loneliness. A 2018 study looking at whether smartphone presence (even with no use) affected psychological and physiological responses to social stress showed that smartphones had a positive effect when it comes to facing social exclusion (Hunter et al.). In this study, participants were faced with a social exclusion stressor, but before the stressor, all participants were randomly sorted into three groups: (1) phone use allowed and encouraged, (2) restricted phone use, or (3) no phone access. The result of this experiment was that both groups that had access to a phone (encouraged use and restricted access) reported lower feelings of exclusion compared to the participants with no access to a phone. This study highlights how the presence of a phone, even with no use, can buffer against the stressor of social exclusion. These results are a powerful indication that phones are used to reduce forms of social distress instigated by acute social stressors.

In addition to alleviating negative emotions associated with stressful social experiences, evidence suggests that certain forms of digital social connection may offer long-term social benefits. For example, a 2021 study investigated whether smartphones help or hinder social connections over time through social support (Lapierre and Zhao). The researchers surveyed 221 college students twice over three months, examining different types of smartphone use: person-to-person, social networking, and mass-mediated. The study found that person-to-person smartphone use correlated with increased feelings of belonging (i.e., feeling accepted by others)

and tangible support (i.e., believing one can find help for practical needs) over the three months. However, the researchers cautioned that mass messaging and social networking did not yield similar benefits. Moreover, problematic smartphone use was associated with decreased tangible support after three months. These findings highlight that while certain forms of digital connection can be socially beneficial, the advantages depend on the specific type of media use. This nuanced relationship between digital media use and social belonging or connection may explain the contradictory research findings in this field. Different types of media appear to have varying effects on individuals, underscoring the complexity of using digital media as a tool for social support and stress relief.

The COVID-19 pandemic served as a natural experiment because it was characterized by reduced social interaction and led to increased feelings of loneliness for many. An Australian study conducted, during this period of mass social distancing, investigated how socially motivated internet use mediated the relationship between health anxiety, isolation behaviors, and depression (Stuart et al.). In other words, the association between increased health anxiety and higher levels of depression was partially explained by increased socially motivated internet use. The researchers found that in conditions of isolation, online social connections effectively buffered against the negative effects of health anxiety. This finding underscores the potential of digital media as a valuable tool for maintaining social connections and supporting mental health during periods of physical isolation.

Other research has shown more direct evidence of digital media use reducing loneliness. For example, one study from 2012 looked at the psychological effects of posting status updates on Facebook (Deters and Mehl). They had one group of participants who were asked to post on Facebook more than they usually do, and then there was a control group who received no direct instruction on what they had to do. The study found that those who were asked to be more active and post more on Facebook had reduced loneliness. This reduction was tied to the fact that participants felt more connected to the people close to them on a daily basis. Another important note about these results is that people's loneliness is reduced regardless of the social feedback they receive. This study highlights how using digital media and using it at a higher rate can make people feel like they have a stronger community.

Digital Apps For Coping With Psychopathology And Stress

Mobile applications have emerged as a popular tool for improving mental health and overall wellness. In 2023, over 50 million people used wellness apps, generating \$950 million in revenue (Wylie). This widespread adoption indicates that people worldwide are leveraging smartphone technology to attempt to enhance their quality of life. Recent research supports the public's belief in the efficacy of these apps.

A 2017 study investigated the potential of internet-based apps to reduce stress levels among college students (Harrer et al.). The researchers recruited college students with elevated stress levels and divided them into two groups: an intervention group that used a mobile app providing stress management techniques with feedback, and a control group. The results revealed

a significant difference in stress levels between the intervention and control groups. Based on these findings, the researchers concluded that app-based interventions can effectively help college students mitigate the effects of stress.

A 2016 study looked at whether a mobile app called Virtual Hope Box was effective in improving coping skills, reducing suicidal ideation, and increasing perceived reasons for living for participants who are at higher risk of suicidal ideation (Bush et al.). The participants of this study were U.S. service veterans who were receiving mental health treatment and who had showcased suicidal ideation. The Virtual Hope Box app helped the participants with emotional regulation and stress management through content targeted at coping strategy practice. What was found is that the participants using Virtual Hope Box were able to cope with negative emotions and unpleasant thoughts significantly better than the control group.

Another app called BioBase was studied to see if it would help users reduce stress and improve mental well-being (Ponzo et al.). The participants were full-time undergraduate students and they were encouraged to use the app daily and complete one therapeutic course. The study found that the four-week intervention of using the BioBase app led to a significant reduction in anxiety and increased perceived mental well-being. After the four weeks, the participants using the app also experienced a reduction in depression levels. In the two-week follow-up from the first four-week period, the positive effects of using the app remained.

Looking at these three studies suggest that using digital media in the form of apps targeting mental wellness can have positive effects on people. However, one important thing to note is that in all three of these studies, all the participants were described to be in stressed states, so the improvements they faced might not be the same as a person with typical stress levels.

Digital Media As A Distraction From Other Aspects Of Life

While there is some evidence that digital media use can be an effective coping strategy for both negative emotions and social problems, other research suggests this coping strategy can become detrimental when it distracts from other important aspects of life. People are reporting media use as a coping strategy and evidence has shown that this coping strategy may increase with age from childhood to adolescence. Adolescence is a critical time for developing critical adaptive coping strategies (Eschenbeck et al.). As such, if adolescents become reliant on media to help manage their emotions, they may not cultivate stress tolerance or practice other adaptive forms of stress management. In the following sections, we will discuss potential negative impacts on social relationships, academics, and the workplace.

Distractions From In-person Interaction (Negative Impact on Relationships)

Phubbing, created by combining the words “phone” and “snubbing”, is the use of smartphones during in-person interactions. In a study from 2021, researchers wanted to look at the predictors and effects of phubbing (Sun and Samp). Through their research, they were able to find out that people who had higher levels of depression, social anxiety, and neuroticism were more likely to participate in phubbing. This suggests that while not all phubbing is necessarily

due to using media to cope with anxiety revolving around the current social situation, it is possible that some phubbing is a form of maladaptive coping with social stress. They also found that agreeableness was negatively related to friend phubbing. The authors found that friend phubbing led to lower levels of friend satisfaction. This study pinpoints how people who are more uncomfortable in social situations use their phones more often in these situations, leading to friend phubbing. As the study shows, this ultimately hinders friendships. This shows how people using digital media/their phones, to cope with the difficulties of social interaction can ultimately have negative impacts on relationships that matter to them.

Another study wanted to dive deeper into how phubbing relates to loneliness and life satisfaction. This study examined 513 college students and the authors primarily examined friendships and their relationship to phubbing (Sun and Wong). Through online self-report measures, it was shown that loneliness was a positive and indirect predictor of friend phubbing. This means that people who tended to be lonely were more likely to phub with their friends. The authors found that this in turn creates lower life satisfaction. This study highlights that people using their phones to feel more comfortable can actually hurt the relationships they have. This is a contrast to what was seen in the s where use of the phone in difficult social environments can be beneficial.

A recent study from 2022 found that smartphone dependence, a behavioral condition where someone has an uncontrollable use of smartphones that hinders or interrupts daily lifestyle activities like relationships, work, or health (Nawaz) or addiction, was positively related to phubbing (Safdar Bajwa et al.). Addictive use like this often includes feelings of distress when smartphones are not available implying a reliance on the phone to relieve such aversive emotions or experiences. Accumulating work suggests that active social media use (ASMU) may share some of the same characteristics as other addictive disorders such as use for mood modification (Goldberg, 1996). This suggests how people who might be most reliant on digital media (possibly as a coping tool for dealing with stress) might be more likely to phub. Further, as multiple studies show, phubbing will lower friendship satisfaction (Sindisari Parus et al; Sun and Yoon).

Family Phubbing

Phubbing does not just happen in friendships; it can also happen in familial settings. One study published in 2020 wanted to look in particular at how parental phubbing affects kids (Pancani et al.). They found children felt more socially disconnected from their parents when they reported a higher level of instances in which they observed their parenting phubbing. This study can help make the case that parents who might be using their phones as a means to cope with stressors in their lives may ignore their children or phub them. This ultimately contributed to children feeling more distant from their parents.

Phubbing also can hinder the relationship between partners. In a study from 2021, researchers were able to confirm that partner phubbing is negatively related to relationship satisfaction (Beukeboom and Pollmann), similar to findings regarding friendships and

parent-child relationships. However, in this study, Beukeboom and Pollmann also found that shared phone use moderates the negative effects of phubbing. Specifically, if partners let each other know what they are doing on their phones, phubbing is not as harmful to the relationship. This can suggest that if someone is using their phone as a coping mechanism it is beneficial to let their partner know what they are doing so that if one partner is phubbing, the other partner knows their intentions. This finding brings nuance to phubbing outcomes because it highlights how it is possible that with better communication, phubbing might not be as detrimental to relationships. More research is needed to know if the same phone-sharing protecting effect exists in parent-child and friendship phubbing situations.

Distractions from Academics/Workplace

The increase in smartphone use over the past few decades has led smartphones and other digital devices to be increasingly present and used in places like academic institutions and the workplace. As the workplace and educational environments may be particularly stressful environments, for those who use their smartphones to cope, smartphone use in these settings may lead to procrastination and possibly more stress down the line. In fact, a study in 2019 examined the effects that digital distractions had in meetings (Malan). Meetings with digital distractions, like smartphones, that were available and used were related to an increase in cognitive load and higher attention conflicts than when digital distractions were not present. It was also found that the meetings with the digital distractions were less productive. They found that meetings with digital distractions took 30% longer to make the same amount of decisions as a meeting with no digital distractions. These results suggest that if people are using digital media in the workplace, it can ultimately make people more distracted and less productive.

More work is needed to understand the possible emotional benefits of media-related coping in these environments and whether they outweigh potential costs to productivity. Further, in an academic setting, a 2012 study examining college students revealed that over half of the participants perceived instant messaging as detrimental to their academic performance (Junco and Cotten). The research found a positive correlation between higher levels of instant messaging use and students reporting difficulties in completing their schoolwork. Another study from 2009 wanted to look at the cognitive effects of using instant messaging while performing some other task (Fox et al.). Participants were asked to complete a reading comprehension task; one group was uninterrupted and the other had to carry out an instant messaging conversation. They found that groups that instant messaged during the task had a lower efficiency and that the more time a participant spent instant messaging was related to lower reading comprehension scores. The findings from both studies underscore the potential negative impact of instant messaging on students' academic achievement. Interestingly, this result contrasts with the previously discussed benefits of instant messaging for stress reduction (Dolev-Cohen and Barak). This suggests a complex relationship where students may use instant messaging as a coping mechanism for stress, but this very use might contribute to academic challenges. This dichotomy

highlights how using instant messaging as a coping tool can have unintended negative consequences in other areas of one's life, particularly in academic performance.

This study exemplifies the nuanced nature of digital media use as a coping strategy, demonstrating that while it may provide short-term stress relief, it can simultaneously hinder long-term goals and responsibilities. Adding on, there has been extensive research on how media multitasking (i.e., using multiple forms of media at the same time) has multiple negative effects in the classroom (Kostić and Randelović; Matthews et al; May and Elder). Research shows how media multitasking leads to distraction because it interferes with attention and working memory. Media multitasking is also shown to have negative effects on reading comprehension, GPA, efficiency, and student-teacher relationships (Kostić and Randelović; May and Elder). Students switching between different digital media may be trying to relieve the stress or anxiety they feel at school. However, this can ultimately hurt their progression regarding education.

One study from 2021 looked at the connection between “excessive social media use and distraction from an information engagement perspective” (Xie et al.). After questioning 743 Chinese college students, researchers found that excessive social media use may impact an individual's general tendency toward distraction (Xie et al.). This study found regular smartphone bloggers were less able to suppress multiple forms of distractions. This highlights how digital media use at higher levels does not just negatively harm aspects of life that are actively getting displaced but can impact your ability to concentrate on all aspects of your life. Following this logic, the use of digital media as a coping tool might also lead to excessive technology use and in turn, distraction from important parts of life. This is why further research must look at whether these negative impacts of excessive media use for distraction from aversive experiences or stress also apply to active coping strategies leveraging online interventions targeted at mental health.

Conclusion

This paper has showcased the benefits and drawbacks of using digital media as a coping strategy. Emerging research shows that digital media can alleviate negative emotions and provide acute stress relief, but findings also qualify that individual factors and content type can impact how beneficial digital media is as a coping mechanism. As we outlined above, the current literature suggests some specific benefits of digital media use for negative emotion relief such as reduced loneliness, and increased social connectedness through things like social media and instant messaging. Further, mobile apps designed specifically to target mental health and aid in coping were effective in improving mental well-being and stress mitigation, especially for people who have high levels of depression or stress. However, there are also growing findings that indicate using digital media as a coping strategy could lead to distractions in various aspects of life. One such area was social settings, where coping with digital media might lead to phubbing. When people phub their family and their partners, research shows that it generally tends to have negative effects on those relationships and feelings of friendship satisfaction. In addition to distracting from in-person social interactions, using digital media also can lead to distractions in

the workplace or academic settings, which leads to multiple negative outcomes. By outlining both these positive and negative impacts of using digital media to cope, this paper provides a more nuanced understanding of the effects that coping with digital media can have. However, research in this area is still in its infancy.

More research is needed that specifically looks at people who are using digital media for coping. Future research should create distinctions between content and/or the type of digital media that people use when they are stressed. This will allow for insight as to what is beneficial for people and what causes negative effects. Future research should also focus on the duration at which digital media is used when someone is coping so that it is known if there is a certain period of time that people should use digital media if they are using it as a coping strategy. Along these same lines, employing both objective behavior tracking and longitudinal designs could help better characterize the frequency in which individuals are employing digital media coping strategies and the longer-term escalation and impacts of this behavior. Understanding these specifics of use is important because people should use digital media to cope if it is helpful as one could argue it is easily accessible, relatively inexpensive, and doesn't result in as immediate negative health consequences as other maladaptive coping strategies such as addictive drugs, engaging in self-harm or other risk-taking behaviors. Nonetheless, it should not get in the way of daily tasks and activities since the disruption can lead to unintended harm in other areas that can ultimately lead to more stress. Overall, there is a need for increased research so digital media users and mental health care professionals can be more informed when turning to digital media in stressful times.

Works Cited

- Beukeboom, Camiel J., and Monique Pollmann. "Partner Phubbing: Why Using Your Phone during Interactions with Your Partner Can Be Detrimental for Your Relationship." *Computers in Human Behavior*, vol. 124, no. 106932, Nov. 2021, p. 106932, <https://doi.org/10.1016/j.chb.2021.106932>.
- Bush, Nigel E., et al. "A Virtual Hope Box: Randomized Controlled Trial of a Smartphone App for Emotional Regulation and Coping with Distress." *Psychiatric Services*, vol. 68, no. 4, Apr. 2017, pp. 330–336, <https://doi.org/10.1176/appi.ps.201600283>.
- Carver, Charles S., et al. "Assessing Coping Strategies: A Theoretically Based Approach." *Journal of Personality and Social Psychology*, vol. 56, no. 2, 1989, pp. 267–283, <https://doi.org/10.1037//0022-3514.56.2.267>. Accessed 11 June 2019.
- Clark, Jenna L., et al. "Social Network Sites and Well-Being: The Role of Social Connection." *Current Directions in Psychological Science*, vol. 27, no. 1, 19 Dec. 2018, pp. 32–37, journals.sagepub.com/doi/full/10.1177/0963721417730833, <https://doi.org/10.1177/0963721417730833>.
- Deters, Fenne große, and Matthias R. Mehl. "Does Posting Facebook Status Updates Increase or Decrease Loneliness? An Online Social Networking Experiment." *Social Psychological and Personality Science*, vol. 4, no. 5, 20 Dec. 2012, pp. 579–586, <https://doi.org/10.1177/1948550612469233>. Accessed 15 Apr. 2019.
- Dolev-Cohen, Michal, and Azy Barak. "Adolescents' Use of Instant Messaging as a Means of Emotional Relief." *Computers in Human Behavior*, vol. 29, no. 1, Jan. 2013, pp. 58–63, www.sciencedirect.com/science/article/pii/S0747563212002191, <https://doi.org/10.1016/j.chb.2012.07.016>.
- Eschenbeck, Heike, et al. "Development of Coping Strategies from Childhood to Adolescence." *European Journal of Health Psychology*, vol. 25, no. 1, Jan. 2018, pp. 18–30, <https://doi.org/10.1027/2512-8442/a000005>. Accessed 28 Nov. 2020.
- Folkman, Susan. "Stress, Coping, and Hope." *Psycho-Oncology*, vol. 19, no. 9, 26 Aug. 2010, pp. 901–908, <https://doi.org/10.1002/pon.1836>. Accessed 12 Apr. 2019.
- Fox, Annie Beth, et al. "Distractions, Distractions: Does Instant Messaging Affect College Students' Performance on a Concurrent Reading Comprehension Task?" *CyberPsychology & Behavior*, vol. 12, no. 1, Feb. 2009, pp. 51–53, <https://doi.org/10.1089/cpb.2008.0107>. Accessed 15 Jan. 2019.
- Guttmann, A. "Time Spent with Digital Media in the U.S. 2022." *Statista*, 29 Aug. 2023, www.statista.com/statistics/262340/daily-time-spent-with-digital-media-according-to-us-consumers/.
- Harrer, Mathias, et al. "Effectiveness of an Internet- and App-Based Intervention for College Students with Elevated Stress: Randomized Controlled Trial." *Journal of Medical Internet Research*, vol. 20, no. 4, 23 Apr. 2018, p. e136, www.jmir.org/2018/4/e136/, <https://doi.org/10.2196/jmir.9293>.

- Harvie, Helen M. K., et al. "Induction of Acute Stress through an Internet-Delivered Trier Social Stress Test as Assessed by Photoplethysmography on a Smartphone." *Stress*, vol. 24, no. 6, 2 Nov. 2021, pp. 1023–1032, <https://doi.org/10.1080/10253890.2021.1995714>. Accessed 1 Feb. 2022.
- Hunter, John F., et al. "The Use of Smartphones as a Digital Security Blanket." *Psychosomatic Medicine*, vol. 80, no. 4, Mar. 2018, p. 1, <https://doi.org/10.1097/psy.0000000000000568>.
- Johnshoy, Quinn, et al. "Social Media Use Following Exposure to an Acute Stressor Facilitates Recovery from the Stress Response." *Physiology & Behavior*, vol. 223, no. 113012, Sept. 2020, p. 113012, <https://doi.org/10.1016/j.physbeh.2020.113012>. Accessed 27 Mar. 2021.
- Junco, Reynol, and Shelia R. Cotten. "Perceived Academic Effects of Instant Messaging Use." *Computers & Education*, vol. 56, no. 2, Feb. 2011, pp. 370–378, <https://doi.org/10.1016/j.compedu.2010.08.020>. Accessed 20 Feb. 2020.
- Kostić, Jelena Opsenica, and Kristina R. Randelović. "DIGITAL DISTRACTIONS: LEARNING in MULTITASKING ENVIRONMENT." *Psychological Applications and Trends*, 22 Apr. 2022, <https://doi.org/10.36315/2022inpact070>.
- Lapierre, Matthew A., and Pengfei Zhao. "Smartphones and Social Support: Longitudinal Associations between Smartphone Use and Types of Support." *Social Science Computer Review*, vol. 40, no. 3, 3 Feb. 2021, p. 089443932098876, <https://doi.org/10.1177/0894439320988762>. Accessed 19 Feb. 2021.
- Leung, Louis. "Using Tablet in Solitude for Stress Reduction: An Examination of Desire for Aloneness, Leisure Boredom, Tablet Activities, and Location of Use." *Computers in Human Behavior*, vol. 48, July 2015, pp. 382–391, <https://doi.org/10.1016/j.chb.2015.01.068>. Accessed 13 July 2019.
- M. Rus, Holly, and Jitske Tiemensma. "Social Media under the Skin: Facebook Use after Acute Stress Impairs Cortisol Recovery." *Frontiers*, vol. 8, 18 Sept. 2017, www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01609/full.
- Malan, Jacob Jacobus. "The Influence of Digital Distraction on Cognitive Load, Attention Conflict and Meeting Productivity." *Open.uct.ac.za*, 2019, open.uct.ac.za/items/5bc49204-609c-45e5-9386-a8e0ff1122b1.
- Matthews, Natasha, et al. "Media-Multitasking and Cognitive Control across the Lifespan." *Scientific Reports*, vol. 12, no. 1, 14 Mar. 2022, <https://doi.org/10.1038/s41598-022-07777-1>.
- May, Kaitlyn E., and Anastasia D. Elder. "Efficient, Helpful, or Distracting? A Literature Review of Media Multitasking in Relation to Academic Performance." *International Journal of Educational Technology in Higher Education*, vol. 15, no. 1, 27 Feb. 2018, pp. 1–17, link.springer.com/article/10.1186/s41239-018-0096-z, <https://doi.org/10.1186/s41239-018-0096-z>.
- Moyer, Melinda Wenner. "Kids as Young as 8 Are Using Social Media More than Ever, Study Finds." *The New York Times*, 24 Mar. 2022, www.nytimes.com/2022/03/24/well/family/child-social-media-use.html.

- Nawaz, Saqib. "Rethinking Classifications and Metrics for Problematic Smartphone Use and Dependence: Addressing the Call for Reassessment." *Computers in Human Behavior Reports*, vol. 12, no. 100327, 1 Dec. 2023, p. 100327, www.sciencedirect.com/science/article/pii/S245195882300060X, <https://doi.org/10.1016/j.chbr.2023.100327>.
- Pancani, Luca, et al. "'Mom, Dad, Look at Me': The Development of the Parental Phubbing Scale." *Journal of Social and Personal Relationships*, vol. 38, no. 2, 12 Oct. 2020, p. 026540752096486, <https://doi.org/10.1177/0265407520964866>.
- Ponzo, Sonia, et al. "Efficacy of the Digital Therapeutic Mobile App 'BioBase' to Reduce Stress and Improve Mental Wellbeing among University Students: A Randomized Controlled Trial (Preprint)." *JMIR MHealth and UHealth*, vol. 8, no. 4, 11 Jan. 2020, <https://doi.org/10.2196/17767>.
- Rus, Holly M., and Jitske Tiemensma. "Social Media as a Shield: Facebook Buffers Acute Stress." *Physiology & Behavior*, vol. 185, Mar. 2018, pp. 46–54, <https://doi.org/10.1016/j.physbeh.2017.12.021>.
- Ryan, Tracii, et al. "How Social Are Social Media? A Review of Online Social Behaviour and Connectedness." *Journal of Relationships Research*, vol. 8, no. 8, 2017, www.cambridge.org/core/journals/journal-of-relationships-research/article/how-social-are-social-media-a-review-of-online-social-behaviour-and-connectedness/5F24EBEC0BC036A5B9AF8D4816F05E2E, <https://doi.org/10.1017/jrr.2017.13>.
- Safdar Bajwa, Ruqia, et al. "Smartphone Addiction and Phubbing Behavior among University Students: A Moderated Mediation Model by Fear of Missing Out, Social Comparison, and Loneliness." *Frontiers in Psychology*, vol. 13, no. 1072551, 6 Jan. 2023, <https://doi.org/10.3389/fpsyg.2022.1072551>.
- Sindisari Parus, Maria, et al. "View of Phubbing Behavior and Quality of Friendship in Faculty of Public Health, Nusa Cendana University." *Undana.ac.id*, Mar. 2021, ejurnal.undana.ac.id/index.php/CJPS/article/view/3031/2442. Accessed 14 Oct. 2024.
- Stuart, Jaimee, et al. "Online Social Connection as a Buffer of Health Anxiety and Isolation during COVID-19." *Cyberpsychology, Behavior, and Social Networking*, vol. 24, no. 8, 18 Feb. 2021, <https://doi.org/10.1089/cyber.2020.0645>.
- Sun, Juhung, and Jennifer A. Samp. "'Phubbing Is Happening to You': Examining Predictors and Effects of Phubbing Behaviour in Friendships." *Behaviour & Information Technology*, vol. 41, no. 12, 17 June 2021, pp. 1–14, <https://doi.org/10.1080/0144929x.2021.1943711>.
- Sun, Juhung, and Norman Wong. "'I Can't Believe I Phubbed up Our Friendship!': Examining Relationships between Loneliness, Problematic Smartphone Use, Friend Phubbing, and Life Satisfaction." *Human Behavior and Emerging Technologies*, 16 Jan. 2024, www.hindawi.com/journals/hbet/2024/5558587/, <https://doi.org/10.1155/2024/5558587>.
- Sun, Juhung, and Doyle Yoon. "Not My Fault to Phub Friends! Individual, Social, and Technological Influences on Phubbing and Its Consequences." *Human Behavior and*

- Emerging Technologies*, vol. 2023, no. 1, 14 Dec. 2023, www.hindawi.com/journals/hbet/2023/4059010/, <https://doi.org/10.1155/2023/4059010>. Accessed 9 Apr. 2024.
- Wallman-Jones, Amie, et al. "Leave the Screen: The Influence of Everyday Behaviors on Self-Reported Interoception." *Biological Psychology*, vol. 181, no. 108600, 1 July 2023, p. 108600, www.sciencedirect.com/science/article/pii/S0301051123001175?ref=pdf_download&fr=RR-2&rr=7f7ae64c38680f4f, <https://doi.org/10.1016/j.biopsycho.2023.108600>.
- Winstone, Lizzy, et al. "Social Media Use and Social Connectedness among Adolescents in the United Kingdom: A Qualitative Exploration of Displacement and Stimulation." *BMC Public Health*, vol. 21, no. 1, 24 Sept. 2021, pp. 1–15, bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-11802-9, <https://doi.org/10.1186/s12889-021-11802-9>.
- Wylie, Louise. "Wellness App Revenue and Usage Statistics (2023)." *Business of Apps*, 16 Aug. 2023, www.businessofapps.com/data/wellness-app-market/.
- Xie, Jia-Qiong, et al. "The Association between Excessive Social Media Use and Distraction: An Eye Movement Tracking Study." *Information & Management*, vol. 58, no. 2, Mar. 2021, p. 103415, <https://doi.org/10.1016/j.im.2020.103415>.

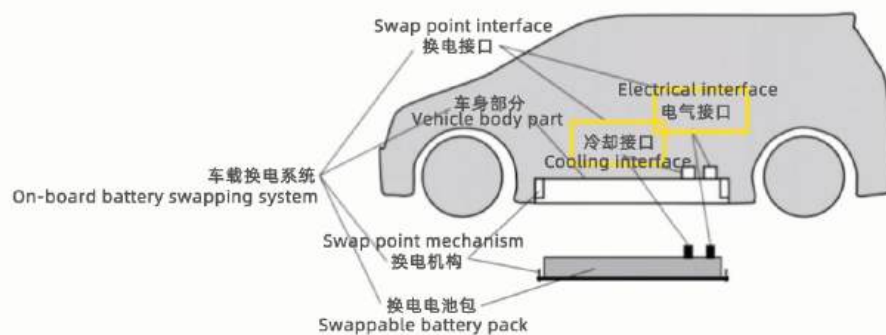
An Analysis of Battery Swap Technology for Electric Vehicles By Xiang Jiang

1. Introduction

In recent years, battery swapping technology for electric vehicles (EVs) has been a continuous topic of discussion in social networks, business reviews, and academia. Chen et al. (2021), for example, stated that both technologically and from a market perspective, this approach is doomed to lead investors into a dead end. Michalek (2021) even called battery swapping a "relic of outdated EV technology"(Page 1). These scholars' criticisms seem to be supported by real-world developments as well. Better Place, one of the early pioneers of battery swapping, went bankrupt in 2013. Tesla, a giant in the EV industry, also discontinued the battery swapping capability for its Model S in 2015. More confusingly, just a month ago, NIO, a leader in battery swapping, revealed in its first quarter financial report that they plan to replace battery swapping technology with 800V ultra-fast charging technology. While all of this seems to suggest against any further development of battery swapping, this article will argue that the advantages of battery swapping in specific scenarios outweigh its drawbacks. The future of battery swapping might not involve one technology replacing the other, but rather the coexistence of both, offering diverse energy replenishment options for EV users.

2. A technical overview

Battery swapping is not a new technique. As early as 1910, battery swapping technology had been put into operation, providing over 6 million miles for the EVs of that time (BBC, 2021). As shown in P1 below, battery swapping involves quickly detaching the vehicle's battery pack from the car body, while using special cooling interfaces to connect the battery pack to the vehicle's cooling system (National standard, 2021). Depending on where the battery is swapped, there are mainly three types of swapping technologies: chassis swapping, modular swapping, and side swapping, with chassis swapping being the most mainstream method today, used by major players like NIO and Aulton.



P1. The Battery Swapping System

3. Advantages

3.1 High Efficiency in Energy Replenishment

One of the greatest advantages of battery swapping is its high efficiency in replenishing energy. As shown in the table comparing four mainstream energy replenishment methods (P2), while a standard gas station and NIO's second-generation battery swap station take roughly the same time to recharge (4 minutes vs. 5–6 minutes), NIO's second-generation battery swap station requires significantly less space. Furthermore, although Aulton's New Energy 4.0 swap station occupies slightly more space (150 square meters), it takes only 1 minute per swap and can serve up to 300 cars per hour, which is more than double the capacity of a gas station or a Tesla V3 supercharging station (a typical example of supercharging stations). Overall, both NIO's and Aulton's battery swap stations have significantly better hourly service capacities than standard gas stations and Tesla's Supercharge stations. Therefore, battery swapping offers advantages in reducing energy replenishment time for users while requiring less space for the same service capacity, which is beneficial for businesses and urban planners.

	Land Occupation	Recharging speed	(Equivalent to the business area of a standard gas station) Hourly service capacity
Standard Gas Station	520 square meters (business area)	4 minutes/ vehicle/filling station	120 vehicles/hour
Tesla V3 Supercharge	12 square meters (land area)	0-80%/32 minutes/parking space	93 vehicles/hour
NIO's second-generation battery swapping station	60 square meters (land area)	5-6 minutes/vehicle/swap station	144 vehicles/hour
Aulton's New Energy 4.0 swapping station	150 square meters (land area)	1 minute/vehicle/battery swapping station	300 vehicles/hour

P2. A comparison of mainstream energy replenishment methods

3.2 Customised Battery Packs and Upgrading Services

Since battery swapping allows for the quick replacement of battery packs, independent users can freely choose different capacities based on their needs, while manufacturers can independently upgrade battery technologies to improve energy density and safety.

Taking independent users as an example, Battery swapping technology allows independent users to choose smaller battery packs, swapping for larger ones only when necessary for long-distance trips. For example, according to NIO's business report (year), in China, its users typically use a 70 kWh battery pack for daily driving and swap to a 100 kWh pack for

long-distance travel. CATL's "Chocolate" battery system is even more flexible, with each module providing around 200 kilometers of range, and up to three modules can be installed to offer around 600 kilometers of range. Battery swapping technology also shows continuous improvement. For example, since 2018, NIO has successively launched 70 kWh, 84 kWh, 100 kWh, and 150 kWh battery packs, with ongoing improvements in structure, thermal management, and battery management systems (citation). This continued development allows battery swapping users to enjoy the benefits of the latest battery technologies earlier than traditional car owners.

3.3 Extended Battery Life

A further advantage is the extended battery life provided by battery swapping technique's centralised battery management. This is because compared to having individual users charging their batteries, having centralised battery stations can more precisely control the battery's charging temperature and speed. According to Yang and Wang (2018), the optimal temperature for charging a battery at maximum efficiency while extending its lifespan is around 30°C. Under constant temperatures, the faster the charging speed, the shorter the battery life. For example, when charging speed increases from 1.5C to 2.5C, battery life drops from 1000 cycles to 500 cycles. By optimizing these charging conditions, swapping stations can at least double the lifespan of batteries. Although regular users might not be highly sensitive to these improvements in battery life, for battery producing companies and society as a whole, extended battery life brings significant economic and environmental benefits.

3.4 Grid Load Balancing

With the increased proportion of renewable energy sources, the fluctuation in supply and demand on power grids becomes more and more pronounced. Battery swap stations can play a key role in grid load balancing.

There seems to be strong evidence supporting this viewpoint from research in China. According to the State Grid Beijing Electric Power Research Institute (2019), most EV owners charge their cars immediately after returning home, creating a charging peak between 8 PM and 10 PM that coincides with residential electricity peak demand. In contrast, battery swap stations can smooth out power consumption by intelligently scheduling battery charging times, reducing peak load pressure on the grid (Rao et al., 2015).

These arguments are echoed in the European context. For example, according to Vallera (2021), the proportion of supply-demand imbalance in Portugal's power grid under battery swapping is the smallest compared to new energy methods such as charging, hydrogen energy, and overhead cables. If fast charging stations are equipped with energy storage devices, it can effectively reduce the pressure on the power grid from fast charging piles. Currently, the mainstream energy storage capacity of solar-storage window charging stations in the market ranges from 250 to 1,000 kWh. A second-generation NIO battery swap station can accommodate 13 batteries, each with a capacity of 70 to 100 kWh, allowing it to store energy ranging from 910

to 1,300 kWh. This capacity is comparable to that of supercharging stations. The 4.0 battery swap station from Aodong New Energy is equipped with 60 batteries, which can store a total of 3,000 to 3,600 kWh, far exceeding the energy storage level of fast charging stations. Some might argue that traditional charging stations can also achieve the effect of transferring stored energy from one period of time to another, thus achieving peak shaving and valley filling. However, charging stations often require customers to purchase additional energy storage devices. For battery swap stations, the batteries stored inside serve as direct energy storage devices, making them more efficient and cost-effective for users and the grid alike.

4. Disadvantages

4.1 Structural Design Compromises

Recent innovations such as Tesla's Structural Battery and BYD's CTB technology deeply integrate battery packs with the vehicle's body, hence improving the vehicle's efficiency and structural strength. However, the need for easy battery removal in swap stations limits the level of integration, which compromises the vehicle's structural strength and energy efficiency. For instance, the table below (P3) shows a comparison among the performance of representative battery-swapping models ET5 ET7 and non-battery-swapping models Tesla Model 3, BYD Seal and XiaoPeng P7 in terms of weight, range, and energy consumption per 100 kilometers. Although all the models are similar in terms of their price and the use of dual-motor technology, the weight of the ET5 is higher than that of TESLA Model 3 and XiaoPeng P7. More notably, the energy consumption levels of the battery-swapping vehicles are higher: both the ET5 and ET7 have energy consumption exceeding 13.6 kWh per 100 kilometers, while the non-battery-swapping vehicles have energy consumption below 12.6 kWh. While it is not possible to assert with certainty that the observed differences in weight and energy consumption are exclusively attributable to the battery-swapping design, the findings from this comparative analysis suggest persuasive that there is likely to be a compromise in efficiency due to the battery swapping design.

Model (dual motor)	Weight	Battery Life (CLTC)	Power consumption per 100km (kwh)
NIO ET5	75 kwh 2000 kg	75 kwh 550 km 100kwh 700 km 150kwh 1000km	13.6-14.18
NIO ET7	2349-2379 kg	75kwh 480-550km 100kwh 615-705km	13.6-15
Tesla Model 3	1897 kg	82kwh 675km	12.15
BYD seal	2150 kg (LFP)	61.4 kwh 550km 82.5 kwh 650km	11.16-12.69
Xiaopeng P7 625E	1890 kg	77.9 kwh 625km	12.46

P3. A comparison of weight and energy consumption between battery-swapping and non-battery-swapping EVs

It is however worth noting that the compromises associated with battery swapping are generally not substantial. Research indicates that the difference in energy efficiency is typically around 10%, and while there may be slightly elevated production costs, these factors are unlikely to be decisive in influencing consumer choices. Given the relatively minor nature of these disparities, it seems premature for companies to simply dismiss the viability of battery swapping techniques.

4.2 Lack of Standardisation

Another drawback concerns the lack of battery swapping standards. To elaborate, a unified standard can benefit both customers and manufacturers. For example, most fast-charging battery models sold in China are equipped with the national standard charging interface (Department Name, 2015). Even models adhering to American and European standards can be converted to the national standard interface, creating a solid foundation for collaboration among brands. This cooperation can lead to significant advantages. For instance, in 2019, Xiaopeng and NIO shared their fast-charging piles, allowing users to charge their vehicles more conveniently and efficiently. Similarly, in 2024, Tesla opened its fast-charging stations to other brands in Europe, further promoting accessibility. These collaborative efforts not only enhance the convenience of charging for users but also increase the profitability of fast charging stations by maximising their utilisation across various brands.

In contrast, the standards for battery swapping technology have been developed much later. There was no official standard until 2020, when the Chinese National Energy Administration released five sets of industry standards for the battery swapping field. In 2021, the Chinese EV Association published another set of standards, outlining recommendations for the development plans of standardised battery swapping stations, as well as technical requirements for swapping platforms and devices, promotion agreements, and recognition systems. Although these two sets of standards represent significant advancements in the field of battery swapping, they are currently treated as voluntary, with no mandatory mechanisms to enforce their implementation. This lack of enforcement risks further development of the battery swapping industry. For instance, NIO's battery swapping stations only support five NIO models, which severely limits their accessibility for users of other brands.

While some may contend that industry standards for battery swapping will further develop as businesses recognise their potential benefits, the likelihood of this occurring is small. Companies may resist the establishment of a distinct set of standards for swappable batteries. As discussed in the previous section, when battery packs are integrated into the vehicle's structure, this design choice can lead to reduced manufacturing costs and enhanced energy efficiency. Consequently, manufacturers might prioritise these advantages over the development of a separate standardized battery swapping system. This resistance within the industry poses

significant challenges to the advancement of battery swapping standards, impeding the feasibility of implementing such technology on a wider scale.

Conclusion

In conclusion, this article provides an analysis of the advantages and disadvantages of battery swapping technologies, revealing that both fast charging and battery swapping occupy distinct positions in terms of energy replenishment efficiency, space requirements, and energy storage capabilities (P4). Additionally, the analysis identifies a significant obstacle to realizing the full potential of battery swapping: the absence of standardised protocols across the industry. Therefore, addressing the issue of standardisation should be a top priority. Given that the vehicle industry is still undergoing a gradual technological evolution from fuel to electric power, battery swapping and fast charging should be viewed as complementary options for users, manufacturers, and governments in promoting the adoption of electric vehicles. It seems premature for scholars and companies to completely overlook the potential of battery swapping techniques.

	Battery-Swapping	Fast-Charging
Efficiency in Energy Replenishment	High	Average
Land Occupation	Low	Average
Storage Capacity	High	Low
Upfront Cost	High	Low
Prevalence	Low	High
Difficulty to profit	High	Low

P4. A Comparison of Fast-Charging and Battery-Swapping

Works Cited

- 21spv.com. (2021). The Energy Administration has approved 7 energy storage standards and 5 battery swap standards, which will take effect in February 2021! - Sunshine Craftsman Photovoltaic Network [online] Available at: <http://www.21spv.com/news/show.php?itemid=78325104132> (2021a). Interpretation of the Standards for the Construction of Battery Swap Stations for Electric Passenger Vehicles, The Energy Bureau has approved seven energy storage standards and five battery swapping standards, effective from February 2021! _ Sunshine Craftsman Photovoltaics [online] People.com.cn. Available at: <http://finance.people.com.cn/n1/2021/1224/c1004-32316247.html>
- Aladdin (2022). NIO Power Has Built This Moat for Four Years - No. 42 Garage [online] Available at: <https://www.42how.com/article/6849>
- Battery100.org. (2024). Design Specifications for Electric Vehicle Battery Replacement (Draft for Public Comments on Partial Revisions). [online] Available at: <https://www.battery100.org/ueditor/php/upload/file/20240528/1716888826394119.docx>
- Battery Swap Revolution2022 | Jiemian News Jiemian · Finance Channel [online] Available at: <https://m.jiemian.com/article/7361824.html>.
- Caam.org.cn. (2016). A Brief Analysis of Electric Vehicle Charging Standards [online] Available at: http://www.caam.org.cn/search/con_5198802.html
- Chen, X., Xing, K., Ni, F., Wu, Y., & Xia, Y. (2021). An electric vehicle battery-swapping system: Concept, architectures, and implementations. *IEEE Intelligent Transportation Systems Magazine*, 14(5), 175-194.
- China-zdgc.com. (2018). Distinctions Between National Standards, Industry Standards, Local Standards, Group Standards, and Enterprise Standards, as well as Their Priority Execution Order - Major Engineering Committee of the China Building Materials Circulation Association, [online] Available at: <https://www.china-zdgc.com/4/9778/126643>
- Desn.com.cn. (2024). Digital Energy Storage Network - Safe Operation and Maintenance - Managed Charging to Ease the Load: How Electric Vehicles Are Shifting from a Burden to an Asset for the Grid. [online] Available at: <https://www.desn.com.cn/news/show-749146.html>
- Diandong.com. (2020). Dialogue with NIO Founder Qin Lihong: Focusing on Digitalization, User Service, Pricing Strategy, and Battery Strategy. [online] Available at: <https://www.diandong.com/news/150850.html>
- Everbright Securities - Carbon Neutrality In-Depth Report (Part Eleven): Charging and Battery Swapping, Collaborative Development for the Range of New Energy Vehicles 210809.pdf https://pdf.dfcfw.com/pdf/H3_AP202108101509209963_1.pdf
- FHWA NHTS BRIEF Electric Vehicle Feasibility <https://nhts.ornl.gov/briefs/EVFeasibility20160701.pdf>

- Gdjn.org.cn. (2021). The Ministry of Industry and Information Technology Promotes Battery Swapping Models! Supporting Pilot Projects in Beijing and Hainan, Vice Minister Xin Guobin Describes Seven Key Advantages of Battery Swapping - Guangdong Energy Conservation Association [online] Available at:
http://www.gdjn.org.cn/hot/show_8008.html
- Haoming Lyu (2020). 2020 National Economic and Social Development Statistical Bulletin of the People's Republic of China - Department of Government Affairs - China Government Website. [online] Wwww.gov.cn Available at:
https://www.gov.cn/xinwen/2021-02/28/content_5589283.htm
- HALLEUX, V. (2021). New EU regulatory framework for batteries: Setting sustainability requirements. [online] Policycommons.net. Available at:
<https://policycommons.net/artifacts/1426784/new-eu-regulatory-framework-for-batteries/2041309/>.
- Jiemia.com. (2022). NIO Achieves Record Monthly Deliveries, Multiple Investment Banks Respond to Details of the Grizzly Bear Short Report. [online] Available at:
<https://m.yicai.com/news/101463168.html>
- McKinsey.com.cn. (2024). A Preliminary Exploration of the Prospects for Battery Swapping in China's Electric Vehicle Industry– McKinsey Greater China. [online] Available at:
<https://www.mckinsey.com.cn/%E4%B8%AD%E5%9B%BD%E7%94%B5%E5%8A%A8%E6%B1%BD%E8%BD%A6%E6%8D%A2%E7%94%B5%E4%B8%9A%E5%8A%A1%E5%89%8D%E6%99%AF%E5%88%9D%E6%8E%A2/>
- Munro Live (2022). Tesla Model Y 4680 Structural Pack OUT!!! [online] YouTube. Available at:
https://www.youtube.com/watch?v=FXpfU6I_T3w
- NIO Inc. (NIO) Q1 2022 Earnings Call Transcript
<https://ir.nio.com/static-files/dde74435-b5a6-4eea-a8b3-13c03a48adae>
- Nea.gov.cn. (2022). National Energy Administration Announcement No. 5 of 2022 - Yunnan Regulatory Office of the National Energy Administration [online] Available at:
https://ynb.nea.gov.cn/xxgk/zcfg/202403/t20240313_258682.html
- Nea.gov.cn. (2024). 11 National and Industry Standards for Electric Vehicle Charging and Battery Swapping Facilities Have Passed Review - National Energy Administration [online] Available at: http://www.nea.gov.cn/2011-12/29/c_131332455.htm
- Netease (2022). XPeng Enters the Battery Swapping Market, Aiming at NIO? [online] 163.com. Available at: <https://www.163.com/dy/article/H0EEE10B051100B9.html>
- Nengyuanjie.net. (2019). Telecharge Officially Reports Profit: Is Spring Arriving for the Charging Pile Industry? - Energy Secto. [online] Available at:
<https://www.nengyuanjie.net/article/26195.html>
- Nio.com. (2021). [online] Available at:
https://app.nio.com/app/web/v2/content/512385?content_type=article&&content_id=512385&_viewBeginTop=true&load_js_bridge=true&show_navigator=false.
- Sohu.com. (2021).

Aulton Builds China's First City-Level Battery Swapping Network in Extreme Cold and Frozen Zones – Changchun. [online] Available at:

https://m.sohu.com/a/501382529_120215473/?pvid=000115_3w_a

Tencent (2023). Exclusive Interview with a Unicorn Company | Yang Ye, Co-founder and Vice Chairman of Aulton New Energy: Pioneering a New Era of Intelligent Battery Swapping _ Tencent News [online] Qq.com. Available at:

https://view.inews.qq.com/k/20231023A09HXR00?web_channel=wap&openApp=false

Ulrich, L. (2021). How Is This A Good Idea?: EV Battery Swapping. [online] IEEE Spectrum. Available at: <https://spectrum.ieee.org/ev-battery-swapping-how-is-this-a-good-idea>.

Vallera, A. M., Nunes, P. M., & Brito, M. C. (2021). Why we need battery swapping technology. *Energy Policy*, 157, 112481.

With thriving energy storage, the layout of integrated 'solar-storage-charging' stations is accelerating. 'Solar-storage-charging' Integrated Stations (2022). “The Energy Storage Landscape in Integrated Charging Stations' - Polaris Power News Network.”[online] Bjx.com.cn. Available at: <https://m.bjx.com.cn/mnews/20220509/1223274.shtml>

Yang, X. G., & Wang, C. Y. (2018). Understanding the trilemma of fast charging, energy density and cycle life of lithium-ion batteries. *Journal of Power Sources*, 402, 489-498.

Yicai.com. (2022). NIO Achieves Record Monthly Deliveries, Multiple Investment Banks Respond to Details of the Grizzly Bear Short Report. [online] Available at: <https://m.yicai.com/news/>

Navigating the Nexus: Tourism, Sustainability, and Homelessness in Hawaii's Changing Landscape By Justin Chun

Abstract

This research explores the nexus between tourism sustainability and homelessness in Hawaii, focusing on how the state's reliance on tourism has contributed to a rising homelessness crisis. Through a survey conducted in August 2024, this study gathers insights into public awareness and opinions regarding the homelessness issue, highlighting a widespread recognition of the problem and its link to Hawaii's booming tourism industry. The main analyses in this paper suggest that the economic benefits of tourism have come at a social cost, with increased housing demand and inflated property prices making it difficult for residents to secure affordable housing. The survey also reveals divergent opinions on the U.S. annexation of Hawaii, with many recognizing its role in boosting infrastructure while others argue that it has exacerbated cultural and social disparities. This study underscores the need for sustainable solutions that balance Hawaii's economic dependency on tourism with the urgent need to address homelessness and promote social equity for its native populations.

Keywords

Homelessness issue, Hawaii, Tourism

Introduction

Homelessness is a pressing global issue that extends beyond individual hardship to encompass significant social, economic, and public health ramifications (Wright, 2000). While those without stable housing often face severe physical and mental health challenges, limited access to essential services, and heightened vulnerability to violence, the broader implications are equally profound. On a national scale, widespread homelessness strains public resources, contributes to increased crime rates, and deteriorates community well-being.

This research explores such dynamics through the lens of homelessness on a remote island, Hawaii, where the issue intersects with tourism, environmental sustainability, and cultural change. Although pollution and cultural disruption are minimal on the island, the presence of homelessness amidst growing tourism highlights how resource demand and environmental pressures can exacerbate social inequalities. Understanding these connections is vital for developing sustainable solutions that address both the immediate needs of homeless individuals and the broader challenges of environmental and cultural preservation. Homelessness is often viewed through the lens of immediate social and economic challenges, yet it is intrinsically connected to broader issues of environmental sustainability and cultural dynamics. This research examines the experiences of homelessness on a remote island, highlighting how such an issue intersects with tourism, pollution, and cultural change.

In the United States, the homelessness issue in some states such as Hawaii, New York, and California are particularly severe due to a combination of high living costs and a shortage of

affordable housing. In Hawaii, the focus of this study, this situation is exacerbated by its geographic isolation that limits available resources and makes living costs even higher. Likewise in California, large cities like Los Angeles and San Francisco have seen significant increases in their homeless populations, creating a visible crisis that affects local communities and puts immense pressure on public services.

The recent surge in homelessness in Hawaii has emerged as a significant issue that impacts the local community and raising concerns about the effectiveness of current policies and interventions. This research aims to explore the public's understanding of the various factors contributing to this growing challenge, including economic, social, and structural influences. By engaging with the community through a comprehensive survey, this study seeks to gather valuable insights into the awareness levels surrounding homelessness in Hawaii and to identify potential solutions that may be effective in addressing this critical issue.

As it turns out, a survey conducted in August 2024 reveals that respondents are in general aware of the problem of Hawaii's homelessness and point to a fact that it is indeed a serious problem. For instance, the U.S. annexation of Hawaii brought its own sets of problems, including the marginalization of native Hawaiians and the erosion of traditional culture, not to mention the homelessness issue. At the same time, many respondents acknowledged that Hawaii island being part of U.S. was a major step to build the infrastructure that boosted Hawaii's economy, without which the islands would have remained vulnerable to bad economic shocks and lack of resources. In that vein, I conclude that while the U.S. annexation of Hawaii introduced significant challenges, it also played a pivotal role in establishing the infrastructure and economic stability that have allowed Hawaii to thrive in the modern world. The complexities of history are undeniable, but it's clear that the benefits of being part of the U.S., including access to resources, trade, and security, have provided Hawaii with the resilience needed to navigate both its unique challenges and opportunities.

This is an important question as understanding the public's perception of homelessness helps to develop informed strategies that can mitigate its impact on the Hawaiian community. This research will not only assess the general awareness of the causes behind the rise in homelessness but also explore public opinion on possible solutions. The insights gathered from this study will be instrumental in shaping future policies and initiatives aimed at reducing homelessness and improving the quality of life for those affected, and this work represents a significant step towards progress in that direction.

Literature Review

Homelessness is a social issue that has been extensively studied across various subjects such as sociology, public health, and urban planning. Research on this matter concludes to a point that homelessness is intertwined with issues such as poverty, mental illness, substance abuse, domestic violence, and systemic inequalities. A study by Kushel et al. (2003), for example, found that homelessness often results from a mix of factors including inadequate access to affordable housing, economic instability, and insufficient social support systems. Moving

further, homelessness issues can be understood at the personal level such as race, gender, and disability. These factors can compound an individual's vulnerability to becoming homeless, as documented by Bassuk et al. (2010).

Homelessness presents itself in different ways across various socio-economic and cultural contexts on an international scale. Although important to note, these differences arise from common underlying factors such as economic inequality, inadequate social safety nets, and housing shortages. In high-income countries such as the United States, Canada, and parts of Europe, homelessness is frequently associated with systemic factors. Factors that contribute to this situation are austerity measures, the reduction of the welfare state, and the demands of urban redevelopment (Benjaminsen & Andrade, 2015). In contrast, in low and middle-income countries, homelessness often results from rapid urbanization, lack of affordable housing, and inadequate infrastructure. For instance, studies by Tipple and Speak (2009) suggest that in cities such as Lagos and Mumbai, homelessness often occurs due to forced evictions and slum clearances aimed at improving the appearance of the urban area. Amore et al. (2013) state that solving homelessness worldwide requires addressing both short-term needs, such as shelter and services, and long-term solutions, such as policy changes to make housing affordable and promote social inclusion. The United Nations' "Habitat for Humanity" and the "Global Homelessness Action Plan" are global efforts to reduce homelessness, but they face challenges due to political and financial limitations.

Narrowing down to Hawaii (the focus of this study), where the rate of homelessness per capita is renowned for being one of the highest in the United States, the situation is worsened by a distinctive blend of economic and geographical factors. Hawaii is infamous for its high cost of living and limited availability of affordable housing. A recent study conducted by Yuan et al. (2018) suggests that the state's heavy dependence on tourism may play a significant role in contributing to this issue. This reliance can drive up housing prices and limit the availability of affordable housing options. Additionally, a considerable portion of Hawaii's homeless population is made up of Native Hawaiians and Pacific Islanders, highlighting deep-seated issues of colonialism, cultural dislocation, and socio-economic inequalities. Local policies like the "Housing First" program are helping to reduce long-term homelessness by giving people stable housing without conditions. However, there are still challenges because of limited funding and resistance from the community. Another recently published paper (Subica et al., 2023) underscores the unique vulnerability of Native Hawaiians and Pacific Islanders (NH/PI) who make up a significant portion of the homeless population. According to this paper, approximately 94% of those surveyed identified as Native Hawaiian, with many facing compounding issues such as unemployment, low educational attainment, and poor physical health. All these factors eventually lead to mental illness of those facing homeless issue, making it difficult for those Native Hawaiians and Pacific Islanders to transition out of homelessness.

Data and Sample Description

I have conducted a survey using SurveyMonkey in August 2024. The survey consisted of two sections. The first section asked about the background information of the respondents. This was done to test if there were differences in people's answers based on age, gender, or education level. I then proceeded to address specific questions about the homelessness issue in Hawaii. The survey was distributed through the SurveyMonkey platform, and I was able to collect responses from a total of 112 respondents. This large number of responses provide a diverse and insightful dataset for analysis. A thorough analysis of the primary demographics and responses is presented below, with a keen emphasis on comprehending the context and the significant implications pertaining to the research. For that reason, I delineate several personal characteristics of respondents in the next subsections.

Data Demographics

Gender Distribution: The respondents were evenly split by gender, with 56 identifying as male (50%), 54 as female (48.21%), and a small minority (1.79%) opting not to disclose their gender. This balance ensures that the findings are not significantly biased towards one gender.

Age Distribution: A significant majority, 45%, were born between 2001 and 2006, indicating that a considerable number of respondents were in their late teens to early twenties, most likely high school or university students. Around 25% were born on or before 1999, while approximately 21% were born in 2000. A very small percentage, around 1%, were born on or after 2007. The overwhelming number of younger participants is intentional and marks the uniqueness of this study, as those respondents are likely to share the similar background with the author of this work who is currently attending a high school in Hawaii.

Nationality Distribution: Over 70% of the people who answered were from the United States, especially those who knew about issues in Hawaii, which is important for the study about Hawaii. The remaining respondents were globally diverse, from regions such as Latin America, Asia, Europe, and Africa. This global perspective, though secondary, adds depth to the findings, particularly when comparing local and global perceptions of the issues.

Other Characteristics of Respondents

In one of the questionnaires, the survey also explored the frequency of travel in 2023-2024. The results revealed that more than 70% of respondents embarked on at least one trip during this period. About 50% reported traveling once or twice, while 25% traveled three or more times. This question was intended to gauge the willingness to travel to other states (countries), which I believe is essential in forming views on challenges related to homelessness. For instance, if you don't travel much and mostly stay in your local area, it can be hard to understand the problem of homelessness. Your familiarity with the area is limited to the places you usually visit, which hinders your full comprehension of the scale of the problem.

Answer Choices		Responses
No, I have never been to Hawaii	45.54%	51
Yes (I do not live in Hawaii, but I have visited the place before)	33.04%	37
Yes (I currently live in Hawaii as a student or as a resident)	19.64%	22
Prefer not to say	1.79%	2
Total		112

Fig 1. The responses to Hawaii visit survey about the following question: *"Have you ever visited Hawaii, an island state of the U.S.?"*

Moreover, the survey placed a significant emphasis on firsthand experiences in Hawaii, which is highly relevant to the research question at hand. This is why I asked the respondents if they had any travel experience to Hawaii. As it turns out, 20% of the people live in Hawaii, while 33% have visited the state. This means that a sufficient 53% of the participants have firsthand exposure to Hawaii, which allows them to provide profound insights into the challenges it faces, particularly the issue of homelessness.

Survey Results and Implications

In this section, I enumerate responses from people who participated in the survey. In particular, there were four questions (one multiple-choice question where respondents were able to select more than one item; the other three were open-ended questions) that were meant to elicit as much as information from survey participants. One of the unique parts of my survey is that I redirected participants to consult to external materials such as Youtube video clip, titled *"Why can't Hawaiians afford to live in Hawaii?"* and short newspaper text, titled *"Federal report: Hawaii has nation's highest rates of chronic homelessness, youth without shelter"*. Take for instance, the questionnaire number 9 (*"What do you believe are the biggest challenges faced by homeless individuals in Hawaii?"*). The responses as recorded in Figure 2 make it clear that the overwhelming majority identified the lack of affordable housing as the primary issue, with 37.5 percent of respondents selecting this option.

Answer Choices		Responses
Lack of affordable housing	64.29%	72
Employment opportunities	48.21%	54

Substance abuse (e.g. use of illegal drugs)	29.46%	33
Mental health services	12.50%	14
Safety and security	16.96%	19
Total		112

Fig2: Responses about the following question:

“What do you believe are the biggest challenges faced by homeless individuals in Hawaii”

When respondents were asked about their priorities as potential governors of Hawaii, the overwhelming majority indicated that addressing the affordable housing crisis would be their primary focus. The issue of affordable housing stems from various factors, one significant contributor being the rising influx of tourists flocking to the island. About 20 percent of those surveyed live in Hawaii, and many believe the rise in homelessness there over the past few years is partly due to increased tourism. This correlation indicates that the surge in visitors could be exerting extra pressure on local resources and housing availability, exacerbating the homelessness crisis. Many believe that the rapid growth in tourism has raised property prices and rental costs, making it harder for locals to find affordable housing. Thus, the interplay between tourism and social issues like homelessness is becoming a growing concern for the residents of Hawaii.

Not everyone recognizes this issue, but I've observed that the island's residents are significantly more attuned to it. The tourists who visit have no clue how their visit damages the local population. Hawaii welcomed about 9 million tourists last year, with around 6 million visiting Oahu. It was recorded that the area of Waikiki had around 80,000 tourists per day. With that many people needing to be accommodated with living spaces, places to eat, etc. The island is getting more expensive by the day, and the local people are negatively affected by the tourists. In 2023, there were about 5,500 homeless people, and this number is rising as housing becomes more unaffordable due to increased tourism.

In response to the final question, *“Do you believe that Hawaii would have been a better place if the United States hadn't annexed the islands? For instance, do you think that the homelessness issue wouldn't have arisen if Hawaii hadn't been part of U.S. territory?”* opinions were divided into two distinct schools of thought. A majority (58% of respondents) believed that Hawaii would not be better off without the U.S. annexation, highlighting both economic and cultural ties as essential factors in their reasoning. These individuals argue that if the United States had not played a role in Hawaii's history, another foreign power might have claimed and even potentially invaded the islands, which could have had far-reaching consequences for the local population. They also suggest that without U.S. support in trade and infrastructure, Hawaii's economy might not have developed as robustly, leaving the islands vulnerable to poverty and lack of resources. Furthermore, they assert that being part of a global power like the

U.S. has offered Hawaii security, technological advancement, and access to global markets—benefits that would have been difficult to achieve in isolation.

On the other hand, a smaller but notable percentage of people argue that U.S. annexation brought its own set of problems, including the marginalization of native Hawaiians and the erosion of traditional culture. These critics claim that the socio-economic challenges seen today, such as homelessness, may stem from the systemic changes imposed by outside influence, which disrupted local ways of life. Nonetheless, I believe no state or country is flawless, and every nation faces its own unique challenges. While Hawaii has certainly encountered difficulties as part of the United States, the benefits—such as economic opportunities, access to modern resources, and security—far outweigh the downsides. Ultimately, Hawaii’s integration into America has helped it grow into a vibrant, culturally rich state that balances its heritage with the advantages of being part of a larger nation.

Conclusion

To conclude, this study highlights the intricate relationship between tourism, homelessness, and cultural sustainability in Hawaii. The influx of tourists has undeniably boosted the state's economy, but it has also exacerbated housing shortages and inflated costs, contributing to the rise in homelessness. To address this issue from young generations' point of view, I have circulated survey questionnaires in August 2024 to collect their reactions to this agenda. As it turns out, survey responses reflect a tendency that is increasingly aware of these issues, particularly the local population. The findings in general suggest that while tourism remains an essential part of Hawaii’s economy, it also creates disparities that disproportionately affect marginalized communities, particularly Native Hawaiians and Pacific Islanders.

Moreover, the research underscores the importance of a balanced approach to addressing homelessness in Hawaii, one that considers both immediate needs, such as affordable housing and social services, and long-term strategies that involve sustainable tourism and cultural preservation. The historical context of U.S. annexation has left lasting impacts on Hawaii’s socio-economic landscape, and the challenges it faces today are rooted in these historical dynamics. Therefore, any effective policy aimed at reducing homelessness must also address the broader issues of inequality and cultural marginalization that have persisted over time.

Works Cited

- Amore, K., Viggers, H., Baker, M.G., & Howden-Chapman, P. (2013). Severe housing deprivation: The problem and its measurement, *Official Statistics Research Series*, 6. Retrieved from: <https://statsnz.contentdm.oclc.org/digital/collection/p20045coll4/id/303/>
- Bassuk, E. L. (2010), Ending child homelessness in America, *American Journal of Orthopsychiatry*, 80(4):496-504
- Benjaminsen, L., & Andrade, S.B. (2015) Testing a Typology of Homelessness across Welfare Regimes: Shelter Use in Denmark and the USA, *Housing Studies*, 30(6):858–876
- Kushel, M. B., Evans, J. L., Perry, S., Robertson, M. J., & Moss, A. R. (2003), No door to lock: victimization among homeless and marginally housed persons, *Arch Intern Med*, 163(20):2492-2499
- News Article: "Federal Report: Hawaii Has Nation's Highest Rates of Chronic Homelessness, Youth Without Shelter." *Hawaii News Now*, 16 Dec. 2023, www.hawaiinewsnow.com/2023/12/16/fed-report-shows-hawaii-has-highest-rates-chronic-homelessness-youth-without-shelter/.
- Subica, A. M., Sampaga, D. S., Ortiz-Misiaszek, J., Martin, T. K. K., & Okamoto, S. K. (2023), The Mental Health, Substance Use, Physical Health, and Mental Health Treatment Need of Community Individuals Experiencing Homelessness in Hawaii, *Community Mental Health Journal*, 59:1021-1026
- Tipple, G., & Speak, S. (2009). *The Hidden Millions: Homelessness in Developing Countries* (1st ed.). Routledge. <https://doi.org/10.4324/9780203883341>
- Wright, T. (2000), Resisting Homelessness: Global, National, and Local Solutions, *Contemporary Sociology*, 29(1):27-32
- Yuan, S., Stern, I. R., Gauci, K. T., & Liu, L. (2018). *Street Youth Study*. Honolulu, HI: University of Hawai'i, Center on the Family.
- YouTube Video: "Why Can't Hawaiians Afford To Live In Hawaii?" YouTube, uploaded by AJ+, 2022, www.youtube.com/watch?v=WZvKsfcmO0M.

Effects of Monetary Sanctions on the Intergenerational Financial Stability of Families with Incarcerated Members By Hyeju Yoon

Abstract

Monetary sanctions are one of the most pervasive and pernicious forms of punishment in the U.S. criminal justice system. The assertion that these sanctions demonstrate justice through retribution, deterrence, and restoration is credible. Yet, they pose an immeasurable challenge to those involved with the justice system, such as the convicted individual and their families. The detrimental effect of monetary sanctions is crucial to reevaluate as it is indicated that their influence can last generations, altering financial stability. By exploring three existing gaps in the current literature regarding how monetary sanctions affect intergenerational families with incarcerated members, this paper aims to provide a legitimate explanation as to the importance of longitudinal studies that capture the effect of financial strain that is heavily tied to not only stability but also making wise decisions concerning finance. Additionally, it focuses on the need for increased awareness of statewide policies and laws in order that modifications can be made to extend them nationwide to mitigate the hardships, thus abating poverty.

Introduction

Monetary sanctions are a deeply embedded element of the U.S. criminal justice system, orchestrated by both federal and state laws. These sanctions trace back to ancient times when such punitive practices were found in legal systems like the Magna Carta, affirming their role in the justice system (Martin et al., 2018). Although they were initially created to serve justice through retribution, deterrence, and restoration, legal punishments often disproportionately burden impoverished families and communities of color. Worse, they extend beyond the short-term impact to prolonged consequences, destabilizing the financial stability of families for generations.

This paper discusses the effects of monetary sanctions on the intergenerational financial stability of families with incarcerated members, a critically unexplored area. Despite the increased literature on the immediate economic effects, a significant gap exists in understanding the transferred generational responsibilities. Specifically, this study addresses the need for longitudinal research that tracks families over time, the psychological and social consequences of monetary sanctions on financial decision-making, and the inadequacies of spreading awareness of the current policies and legal frameworks in mitigating these burdens.

By analyzing these gaps, this paper aims to understand better how legal debt perpetuates cycles of poverty and financial instability across generations. In addition, it contributes to the ongoing controversies regarding the need for more equitable and effective legal and policy interventions. It highlights the urgent necessity for policymakers and advocates to rethink the role of monetary sanctions in ensuring not only justice but also equality and economic well-being for the families affected.

Methods

A literature review was conducted using Google Scholar. Keywords including “intergenerational financial stability” and “monetary sanctions” were initially used to find articles that would discuss the broad topic of how monetary sanctions and intergenerational families are interconnected. Then, specific keywords were used, such as “policies and laws” and “psychological influence,” to narrow the topic and identify key gaps in the current literature. As many studies as possible were accumulated and analyzed to provide legitimacy for the paper, allowing for an extensive discussion on this specific topic.

Literature Review

Longitudinal Studies on Intergenerational Effects

Within the criminal justice system, monetary sanctions and other financial penalties are pervasive forms of punishment. Unlike community service, probation, or incarceration, legal financial obligations (LFOs) can be passed on to family members. Research has increasingly highlighted how these sanctions disproportionately affect marginalized families, exacerbating financial difficulties and emotional stress, particularly those of color, thus reflecting broader racial and class disparities in the justice system.

Additionally, such burdens are amplified by additional costs of poverty, also known as the “poverty penalty,” including compound interest, late fees, and prolonged supervision fees, making it nearly impossible for affected families to recover financially (Horowitz et al., 2024). The financial strain is eminent in cases where defendants in different states face varying monetary penalties for similar offenses (Horowitz et al., 2024). In California, a defendant may face sanctions ranging from \$1,310 to \$3,480 for a first-time driving offense with a suspended license. In contrast, a similar defendant in North Carolina or Minnesota may expect a much alleviated penalty of \$188-\$388 and \$279.50-\$289.50, respectively. Despite these geographical disparities, such a burden on families who support their loved ones remains a complex problem, typically leading them to deplete savings and pawn valuable items to cover the costs (Boches et al., 2022).

Beyond these immediate financial burdens, the broader impact of monetary sanctions extends to multiple generations of families. Research such as that of Horowitz et al. (2024) indicates the perpetuation of poverty, exploring the trigenerational impact of monetary sanctions in Minnesota. Their findings reveal that the fiscal burdens placed on parents often lead to significant stress and resentment as they focus on providing aid to support their adult children’s legal expenses. For example, Lisa and Sharon, interviewees in the study, noted that their parents struggled to manage their financial responsibilities due to the costs associated with their children’s legal issues, such as court fines and attorney fees. Children are equally affected by the economic constraints imposed by sanctions. Horowitz et al. (2024) highlight how parents confronting legal debt often struggle to provide for their children’s basic needs, such as food,

housing, and transportation, creating long-term consequences for their development and future stability.

Despite the increase in research on the financial impacts of monetary sanctions, there is a significant gap in studies that track these effects over multiple generations. Most existing research, including that of Halladay (2019), Boches et al. (2022), and Horowitz et al. (2024), focuses on the short-term effects. However, there is a critical need for longitudinal studies that follow families over time to understand how monetary sanctions cause indefinite poverty and economic instability. While Horowitz et al. (2024) begin to address this gap by analyzing trigenerational effects, their study is geographically limited to Minnesota, lacking the comprehensive analysis required to generalize these findings across different states and demographic groups. Moreover, the qualitative nature of their study, while valuable, indicates the need for more quantitative research for legitimate data on the long-term financial pathway of families affected by legal debt.

Psychological and Social Impact on Financial Decisions

The link between psychology and financial decisions is underscored by the strain and deprivation experienced during formative years, which can have lasting detrimental effects on economic behavior. The incarceration of a family member, especially a parent, plays a significant role in shaping the financial decision-making of children as they grow older, particularly within impoverished households. In such environments, persistent economic stress not only destabilizes fiscal stability but also affects the psychological and social development of the children. Existing research, however, frequently overlooks the effects and consequences of having an incarcerated member influence financial behavior and responsibilities in adulthood despite its transferable nature, creating a complex web of economic strain. This may alter children's outlook on money and monetary planning from an early age, playing a critical role in shaping how they make decisions when they come of age (Ashraf & Farhad, 2022; Luk et al., 2022).

Cognitive-behavioral theories further emphasize how the environment shapes behavior, highlighting how children learn antisocial and maladaptive behaviors from observing their parents' interactions with the criminal justice system. Psychological theories, such as cognitive-behavioral theory, suggest that financial behaviors are learned and reinforced within the family environment. Children exposed to economic hardships due to a parent's incarceration may adopt a shortfalling mindset, leading to behaviors characterized by short-term fiscal decisions and aversion to long-term planning (Platt & Prout, 1987). Such behavioral patterns are further complicated by the emotional toll of ongoing monetary instability, which can cause children to prioritize immediate financial needs over future security, often perpetuating a cycle of poverty. The chronic stress associated with economic deprivation interrupts cognitive functions, impairing the ability to make sound financial decisions that induce insecurity even in adulthood.

Additionally, research indicates that parental incarceration often imposes severe economic deprivation, the loss of parental socialization, and the stigma associated with societal labeling (Levin & Topalli, 2019). Children affected by the incarceration of a parent are more

likely to experience negative emotional and behavioral outcomes, including depression, anxiety, and problems with school attendance, academics, and peer relationships. This psychological burden is compounded by the stigma associated with incarceration, which can lead to social isolation and a lack of supportive relationships, further impeding financial stability. Studies have shown that children who experience the imprisonment of parents internalize these experiences, leading to long-term issues in emotional regulations and decision-making that later become discernible in financial milieus. (Turney, 2014).

The chronic stress associated with parental incarceration has particularly hostile effects on children's cognitive skills and overall well-being (Turney, 2014). Incarcerated parents often belong to already vulnerable families, and their absence exacerbates existing economic and social disadvantages. Children in these circumstances frequently face unstable home environments, economic hardships, and poor access to quality education and healthcare, all of which translate to impaired cognitive development and financial literacy. As these children grow older, they are more likely to struggle with making deliberate financial decisions due to this protraction.

Regardless of the apparent psychological and social effects of parental incarceration on children's financial decision-making, there is a notable gap in the literature that directly addresses this issue. Most existing studies focus on the immediate emotional and behavioral impacts of parental incarceration without examining how these factors translate into financial behavior as children grow older. Further research is needed to explore how these early psychological challenges influence long-term fiscal stability and decision-making processes in the children of parents involved with the justice system. Understanding this correlation is crucial for developing interventions that address not only the economic but also the psychological needs of children affected by parental incarceration, ultimately breaking the cycle of disadvantage and fostering more positive financial outcomes for intergenerational families.

Policy and Law and Their Effectiveness

As mentioned earlier, monetary sanctions are among the most pervasive forms of punishment by the criminal justice system in the U.S., serving as a tool for retribution, deterrence, restoration, and revenue generation (Martin et al., 2018). However, while these sanctions serve several functions within the system, their implementation often places a financial burden on those least able to afford it. Despite the recognized financial strain these sanctions place on individuals and their families, existing literature primarily focuses on enforcing these penalties, failing to address equivalent laws or policies to mitigate their impact on affected families. This section explores the legal frameworks and policies implementing monetary sanctions and examines the critical gap in alleviating measures for those burdened by these financial obligations.

In 1966, the Mandatory Victim Restitution Act (MVRA) made restitution mandatory for certain crimes, removing judicial discretion and enforcing restitution regardless of the offender's financial capability (Chase, 2001). This approach underscores the punitive nature of monetary

sanctions, as it compels financially disadvantaged offenders to compensate victims, often at significant personal cost. Additionally, courts impose various fees intended to cover the costs of the justice system, such as fees for public defenders, processing payments, or even daily incarceration costs. State laws set these fees and are often unassociated with the actual expenses imposed by the justice system, further entrenching economic inequities within the penal framework (Martin et al., 2018). The cumulative effect of these sanctions can be overwhelming, with total fees often exceeding thousands of dollars, an insuperable amount that is highly impractical to pay for many offenders and their families. For instance, Harris (2016) found that individuals in Washington faced average legal financial obligations of nearly \$1,300 per felony conviction, with some paying over \$10,000, creating long-term economic hardship and resulting in extended non-payment consequences (Harris, 2016, xxii).

The enforcement of these monetary sanctions extends beyond the courtroom, as law enforcement agencies become debt collectors. Financial pressures on public institutions, including courts and police departments, incentivize the accumulation of fines and fees. For instance, in California, fines and fees are redirected to the General Fund, the primary pool of government revenue used to finance a broad range of public services. Similarly, Nevada's courts and Texas probation departments rely heavily on revenue from these sanctions as well. Law enforcement officers are often tasked with enforcing warrants for nonpayment or focusing on infractions that generate income rather than prioritizing public safety. Such practices can lead to increased interactions between police and low-income individuals, perpetuating a cycle of criminal justice involvement due to the inability to pay. Moreover, states like Texas have implemented programs that allow officers to use credit card readers to collect fines on the spot, presenting individuals with the choice of immediate payment or facing arrest. This creates an inequitable situation, particularly for those unable to afford immediate payment, further highlighting the lack of measures to mitigate financial strain on affected families (Martin et al., 2018).

Collateral consequences of unpaid monetary sanctions, such as driver's license suspensions, disenfranchisement, and potential incarceration, further entrench individuals within the criminal justice system. License suspensions, commonly used across states for nonpayment of fines, significantly impair individuals' ability to work and care for their families, disproportionately affecting marginalized communities. Incarceration for failure to pay, despite being deemed unconstitutional when not willful, continues to occur in practice, reinforcing the punitive approach of the justice system towards financial non-compliance. These practices reveal a systemic reliance on monetary sanctions as a revenue source, with little concern for the broader social and economic impacts on individuals and their families (Martin et al., 2018).

In spite of the extensive implementation of such penalties, there is a notable absence of addressing national laws, policies, or guidelines to acknowledge and mitigate their financial impact on intergenerational families. Current legislative frameworks, such as the MVRA and the Prison Litigation Reform Act (PLRA), emphasize enforcement rather than alleviating financial burdens (Chase, 2024; Schlanger, 2008). Reforms in some jurisdictions, such as those in

Ferguson, Missouri, and Biloxi, Mississippi, have emerged through legal challenges and public advocacy. Still, these changes are localized and must be more widespread (Martin et al., 2018). While some states have taken steps to mitigate the impact of monetary sanctions, such as offering installment payments or reducing fines for indigent defendants, these measures are not uniformly applied and do not constitute a comprehensive policy approach to address the financial strain on families.

The enforcement of monetary sanctions in the U.S. criminal justice system imposes substantial financial burdens on individuals and their families. Current laws and policies primarily emphasize the collection of these sanctions, often neglecting the long-term financial impact on offenders and their families. This lack of mitigating measures exacerbates economic disparities and perpetuates cycles of poverty, revealing a significant gap in the legal framework. To address this, policies should shift toward reducing the financial strain of monetary sanctions, ensuring the justice system does not disproportionately penalize those least able to pay. Future research and policy development must prioritize equitable solutions that balance accountability with the economic realities of affected families, fostering a more just and sustainable approach to criminal justice sanctions.

Analysis

Longitudinal Studies on Intergenerational Effects

The literature review suggests that monetary sanctions imposed on individuals extend beyond immediate effects and have prolonged intergenerational consequences on families. Although existing research has explored the short-term effects of these sanctions, there is a significant gap in studies tracking the overall, extended financial impact on multiple generations.

One of the primary arguments supporting this hypothesis is that monetary sanctions often lead to compounding debt. Eckart (2022) and Turner (2019) emphasize how these penalties are a form of punishment that impairs families, particularly those already struggling with poverty or of color. The Washington report addresses that payments for legal debt can last several years and, at worst, extend well beyond the end of formal punishment, penalizing both convicted individuals and their families, including children and future generations (Brown, 2023). For instance, formerly incarcerated individuals from minority groups, including communities of African American and Hispanic families, face higher rates of legal debt relative to their annual income: formerly incarcerated African Americans owe 50% of their yearly income to sanctions.

Additionally, the stories of Lynn Ausley and Jessica Chavez provide compelling evidence of how monetary sanctions impact families across generations. Ausley is a mother and a grandmother whose son, who identifies as Black, was involved in a drive-by shooting and was charged as an adult despite being underage. He faced a million-dollar bail and a 15-year prison

sentence.⁹³ The compounding fees and sanctions devastated the family financially and emotionally, leaving them with thin resources and a sense of despair (Debt Free Justice, 2023). Similarly, Jessica Chavez's story exemplifies the long-term financial toll that monetary sanctions have on families, even after a long time since the initial crime occurred. Chavez's cousin had been involved in the juvenile justice system since he was nine years old, and even decades later, the financial burden persisted.⁹⁴ Long after her grandson's offenses, Chavez's grandmother remained in collections for fines and fees accumulated from his juvenile court cases, illustrating how such sanctions can outlast an individual's life, not only the convicted but also their relatives, with financial obligations passed down through generations (Debt Free Justice, 2023).

These stories highlight the intergenerational nature of the financial strain caused by monetary sanctions, which not only hinder the immediate family members' financial stability but also create long-term consequences for future generations. As families are forced to prioritize paying off criminal debt over investing in opportunities for financial stability and growth, the absence of longitudinal studies tracking these effects across generations represents a significant gap in existing research. The analysis, including the stories, underscores the importance of more comprehensive research on the intergenerational financial impacts of monetary sanctions as well as understanding how these sanctions perpetuate cycles of poverty, particularly in marginalized communities, where the burden of criminal debt is disproportionately heavy.

Psychological and Social Impact on Financial Decisions

Regarding the current section, the literature review implies that criminal debt stemming from monetary sanctions could limit the financial literacy and education of children of convicted parents, thereby making them vulnerable to poor financial decisions. This aligns with the broader understanding of psychological and social impacts on decision-making. Research by Mendel emphasizes that young people's brain development continues until the age of 25, making them particularly susceptible to risky behavior (Mendel, 2023). The abilities tied to psychosocial maturities - impulse control, delayed gratification, and the capacity to weigh consequences - are critical for sound decision-making. When a parent's economic situation is strained by criminal debt, children are deprived of necessary educational resources and guidance, which would otherwise support their psychosocial development.

⁹³ Ausley recounts how the system seemed unconcerned about her family's suffering, as the sanctions extended far beyond the punishment of her son. She described the feeling of worthlessness as a parent, stating that the system made it seem her son's worth was tied only to the fines levied against him, turning him into what she described as a mere "mechanism" of profit for the criminal justice system.

⁹⁴ Despite her best efforts, such as hosting yard sales and selling food from her garage, she could not pay off the debt before her passing, burdening her with the stress and financial strain until her death. Chavez expressed her anger and frustration at a system that seemed created to penalize poor families like hers indefinitely, with legal debt accumulating endlessly.

Mendel's assertion that limited education impedes the maturation process is crucial. Without access to education or proper guidance, children may struggle to develop the skills needed for responsible financial behavior. This scenario is exacerbated when criminal debt forces parents to focus on immediate survival, leaving little room for teaching financial responsibility. A hypothetical case could involve a single mother whose economic resources are depleted by court fees and restitution following her husband's incarceration. Now the sole provider, she may have to prioritize basic survival expenses over her child's educational needs, such as extracurricular activities and tutoring that could help develop his monetary knowledge. As a result, her son is left without the tools to learn responsible habits, leading to impulsive spending and a lack of savings habits as he enters adulthood.

Supporting this, Saunders (2022) highlights how financial stress in childhood has been linked to long-term adverse effects on mental health and development. Children in precarious households may take on additional responsibilities at a young age, which can hinder their academic performance and contribute to developmental delays. The hypothetical example above mirrors this reality - such insecurities at home cause the son to take on a part-time job during high school, which cuts his study time and leaves him unprepared for higher education and financial independence. As a result, he may become more vulnerable to engaging in risky behavior, similar to his father.

The instance of Julissa Soto and her family provides real-world insight into how financial pressures linked to the justice system can ripple through generations. Soto's son's involvement with the system placed an immense budgetary strain on the family, leading her daughter to drop out of school. This likewise reflects the hypothetical situation where financial stress leads to a cycle of poor fiscal decisions that impact children's future, as they lack the necessary guidance to navigate economic challenges (Debt Free Justice, 2023).

Even financial professionals, such as Erin Wood, acknowledge the significant impact that criminal justice-related decisions can have on families. Wood stresses the need for assistance in guiding families through complex choices, as the strain of managing criminal debt often results in long-term instability. Without external support, the future is limited by a lack of education in money management, hindering both money and psychological maturation and perpetuating the cycle of vulnerability to crime and poor financial decisions (Wood, 2022).

Thus, the reciprocal nature of Mendel's argument holds that while his focus is on adolescents convicted of crimes, the psychological and social impacts of criminal debt on children of incarcerated parents also impede fiscal maturity. In both cases, limited education and financial stress hinder psychosocial development, leaving children exposed to the same behaviors that trapped their parents in the justice system. This interplay between economic strain, limited education, and delayed psychological maturity makes a compelling case for addressing the broader implications of monetary sanctions on intergenerational families.

Policy and Law and Their Effectiveness

The literature review of this section suggests that there is an absence of recognition of national policies or laws that mitigate the impact of monetary sanctions. Furthermore, existing policies and regulations do not equally affect all individuals. Instead, it places undue financial stress on families with incarcerated members, particularly those facing economic constraints, highlighting the importance of local laws and their varied effects across jurisdictions. The identified gap is that although some states and localities have taken steps to alleviate the burdens through specific policies, these efforts are not widespread, and almost no literature explores these laws and their effectiveness in reducing intergenerational fiscal instability.

Local efforts to overcome the financial weight of monetary sanctions provide some relief, but their reach and implementation are inconsistent across the nation. For instance, House Bill 1315 in Colorado eliminates fees for juveniles and their families, reducing the costs associated with probation, legal representation, medical care, and more (Fines & Fees Justice Center, 2021). Similarly, California Senate Bill 190 abolished administrative fees in juvenile delinquency cases, ending charging for services like electronic monitoring, probation, and drug testing (Fines & Fees Justice Center, 2017). These reforms represent a shift toward reducing the financial harm caused by monetary sanctions, particularly on vulnerable populations.

However, California's approach underscores the limitations of these types of bills. While California Assembly Bill 1869 extended Senate Bill 190 by abolishing additional criminal fees and allocating state funds to offset lost revenue for counties, these measures are confined to the state level (Fines & Fees Justice Center, 2020). The absence of national legislation means that families in other states still bear the financial strain of monetary sanctions. Even within California, many counties have not yet reduced or eliminated fees. This indeed underscores the existence of unstable policies to palliate pre-existing debt or refund for those who had already paid these fees, leaving families with prolonged fiscal burdens from past sanctions.

The issue extends beyond California, with reforms within the state limited in scope and impact. For instance, in Washington, House Bill 1412 allows judges to waive legal financial obligations (LFOs) for individuals who can prove their inability to pay, "thereby increasing opportunities for individuals to reintegrate successfully into the community" (Adriana, 2022). This law, however, is restricted due to key sanctions, such as the Victim Penalty Assessment and DNA collection fees, that remain mandatory. Similarly, San Francisco's elimination of criminal justice fees in 2018 demonstrated a progressive reform, lifting \$32 million in debt from 21,000 residents - primarily benefiting Black and brown communities (Office of the Public Defender, 2020). Yet, like House Bill 1412, San Francisco's reform is a localized effort that lacks the national support necessary to address the systemic issues causing financial instability.

While state-level reforms such as House Bill 1315, California Senate Bill 190, and Washington's House Bill 1412 mark significant progress in reducing the financial strain of monetary sanctions, they remain isolated and lack substantial attention. Despite their steadfast existence, the absence of a cohesive national policy leaves a considerable gap, forcing families nationwide to face heavy economic burdens. Moreover, the limited focus on these state-level initiatives in the literature highlights the need for more comprehensive research and discussion

into how national laws could function as a federal strategy to mitigate the economic impacts on convicted individuals and their intergenerational families. Nevertheless, without prevalent action, monetary sanctions will continue to undermine the financial stability of intergenerational families, deepening existing inequalities.

Conclusion

The findings suggest that monetary sanctions have an inimical consequence on intergenerational families with convicted members. No matter the age, the effect is detrimental, straining not only financial stability but also education, which is highly correlated with making informed decisions regarding finance. However, without a policy or law that is prevalent to the extent that it can alleviate the economic pressure for families across the nation, such hardships will perpetuate, exacerbating deprivation. It limits the resources for the prisoner, who, upon release, will face difficulties in financial recovery due to accumulated debt, transferring the burden to their family members who will have to deprive their lives to recuperate financially.

While the findings are significant, some aspects of the study could be explored further due to constraints, including difficulty in acquiring specific information. For instance, the literature review suggested a more quantitative study that could illustrate the extended effects of monetary sanctions on intergenerational families. However, no such data could be obtained due to a lack of awareness, thus limiting the understanding of the harmful influence of monetary sanctions that could last for generations.

Despite the challenges, further exploration is critical in addressing the harmful nature of monetary sanctions. Although they are a crucial component of the criminal justice system, if modifications are not made, poverty and misery will be perpetuated for not only the convicted individual but also their families, further confining them in a system that will confiscate their wealth, education, and liberty.

Works Cited

- Adriana. (2022, April 5). Governor Inslee signs bill providing LFO relief options to people who can demonstrate an inability to pay - Columbia Legal Services. Columbia Legal Services - Seeking Justice Together.
[https://columbialegal.org/governor-inslee-signs-bill-providing-lfo-relief-options-to-people-who-can-demonstrate-an-inability-to-pay/#:~:text=House%20Bill%20\(HB\)%201412%20gives,reintegrate%20successfully%20into%20the%20community.](https://columbialegal.org/governor-inslee-signs-bill-providing-lfo-relief-options-to-people-who-can-demonstrate-an-inability-to-pay/#:~:text=House%20Bill%20(HB)%201412%20gives,reintegrate%20successfully%20into%20the%20community.)
- Ashraf, I., & Farhad, S. (2022). Children's experiences of parental incarceration. In *The Palgrave handbook of global social problems* (pp. 1-12). Cham: Springer International Publishing.
https://link.springer.com/referenceworkentry/10.1007/978-3-030-68127-2_276-1
- Boches, D. J., Martin, B. T., Giuffre, A., Sanchez, A., Sutherland, A. L., & Shannon, S. K. (2022). Monetary sanctions and symbiotic harms. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 8(2), 98-115. <https://www.rsfjournal.org/content/8/2/98>
- Brown, S. M. (2023, December 20). Sentencing Project Says U.S. Took Strides Towards Reforms, Progress, and Ongoing Challenges. *The Washington Informer*.
<https://www.washingtoninformer.com/mass-incarceration-2023-reforms/>
- Chase, I. J. (2001). Making the criminal pay in cash: The ex post facto implications of the Mandatory Victims Restitution Act of 1996. *The University of Chicago Law Review*, 68(2), 463-489.
<https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=5088&context=uclrev>
- Debt Free Justice. (2023, April 4). Julissa Soto. <https://debtfreejustice.org/stories/julissa-soto>
- Debt Free Justice. (2023, June 1). Lynn Ausley. <https://debtfreejustice.org/stories/lynn-ausley>
- Debt Free Justice. (2023, June 26). Jessica Chavez.
<https://debtfreejustice.org/stories/jessica-chavez>
- Eckart, K. (2022, March 2). Multi-state study of monetary sanctions finds widespread inequities, far-reaching consequences. *UW News*.
<https://www.washington.edu/news/2022/03/02/multi-state-study-of-monetary-sanctions-finds-widespread-inequities-far-reaching-consequences/>
- Fines & Fees Justice Center. (2017, January 26). California SB 190: Juvenile fees. Fines and Fees Justice Center.
<https://finesandfeesjusticecenter.org/articles/california-sb-190-juveniles/#:~:text=Previously%2C%20California%20gave%20counties%20the,Incarceration>
- Fines & Fees Justice Center. (2020, October 1). California AB 1869 criminal fees. Fines and Fees Justice Center.
<https://finesandfeesjusticecenter.org/articles/california-ab-1869-criminal-fees/>
- Fines & Fees Justice Center. (2021, September 27). Colorado HB21-1315: Concerning eliminating certain monetary amounts a juvenile in the justice system is required to pay. Fines and Fees Justice Center.
<https://finesandfeesjusticecenter.org/articles/colorado-hb21-1315-concerning-eliminating-certain-monetary-amounts-a-juvenile-in-the-justice-system-is-required-to-pay/>

- Halladay, J. (2018). The Thirteenth Amendment, prison labor wages, and interrupting the intergenerational cycle of subjugation. *Seattle UL Rev.*, 42, 937.
<https://digitalcommons.law.seattleu.edu/cgi/viewcontent.cgi?article=2592&context=sulr>
- Harris, A. (2016). *A pound of flesh: Monetary sanctions as punishment for the poor*. Russell Sage Foundation.
https://www.google.com/books/edition/A_Pound_of_Flesh/c3MADAAAQBAJ?hl=en&gbpv=1&printsec=frontcover
- Horowitz, V. L., Larson, R. P., Stewart, R., & Uggen, C. (2024). Fines, Fees, and Families: Monetary Sanctions As Stigmatized Intergenerational Exchange. *The Sociological Quarterly*, 1-20.
http://users.soc.umn.edu/~uggen/Horowitz_Stewart_Larson_Uggen_TSQ_24.pdf
- Levine, K. L., & Topalli, V. (2019). Process as intergenerational punishment. *The Legal Process and the Promise of Justice: Studies Inspired by the Work of Malcolm Feeley*, 55-71.
<https://www.cambridge.org/core/books/abs/legal-process-and-the-promise-of-justice/process-as-intergenerational-punishment/9E544E6ABD608F8802641693BB1AB7B1>
- Luk, M. S. K., Hui, C., Tsang, S. K. M., Fung, Y. L., & Chan, C. H. Y. (2023). Physical and psychosocial impacts of parental incarceration on children and adolescents: A systematic review differentiating age of exposure. *Adolescent Research Review*, 8(2), 159-178.
<https://link.springer.com/article/10.1007/s40894-022-00182-9#:~:text=The%20earlier%20the%20exposure%20to,often%20continue%20into%20emerging%20adulthood.>
- Martin, K. D., Sykes, B. L., Shannon, S., Edwards, F., & Harris, A. (2018). Monetary sanctions: Legal financial obligations in US systems of justice. *Annual Review of Criminology*, 1(1), 471-495.
<https://www.annualreviews.org/content/journals/10.1146/annurev-criminol-032317-091915#:~:text=The%20review%20describes%20the%20most,driver's%20license%20suspension%20C%20voting%20restrictions%20C>
- Mendel, R. (2023, March 27). Why youth incarceration fails: An updated review of the evidence. The Sentencing Project.
<https://www.sentencingproject.org/reports/why-youth-incarceration-fails-an-updated-review-of-the-evidence/>
- Office of the Public Defender. (2020, September 24). California Becomes the First State in the Nation to End Collection of Fees in the Criminal Legal System. San Francisco Public Defender's Office.
<https://sftreasurer.org/california-becomes-first-state-nation-end-collection-fees-criminal-legal-system>
- Platt, J. J., & Prout, M. F. (1987). Cognitive-behavioral theory and interventions for crime and delinquency. In *Behavioral approaches to crime and delinquency: A handbook of application, research, and concepts* (pp. 477-497). Boston, MA: Springer US.
<https://www.ojp.gov/ncjrs/virtual-library/abstracts/cognitive-behavioral-theory-and-interventions-crime-and-delinquency>

- Schlanger, M., & Shay, G. (2008). Preserving the rule of law in America's jails and prisons: The case for amending the Prison Litigation Reform Act. *U. Pa. J. Const. L.*, 11, 139.
<https://repository.law.umich.edu/articles/1262/>
- Saunders, J. D. (2022, December 20). Understanding How Financial Issues Affect Your Children and What You Can Do to Help. Kiddie Kredit.
<https://www.kiddiekredit.com/post/understanding-how-financial-issues-affect-your-children-and-what-you-can-do-to-help#:~:text=Open%20Discussion%20of%20Money%20Matters&text=It%20is%20important%20to%20start%20having%20financial%20discussions%20early%20with,to%20contribute%20to%20the%20conversation.>
- Turner, H. (2019, February 6). The Price of Freedom: An Analysis of Monetary Sanctions in the United States. *The Classic Journal*.
<https://theclassicjournal.uga.edu/index.php/2019/02/06/the-price-of-freedom-an-analysis-of-monetary-sanctions-in-the-united-states/>
- Turney, K. (2017). The unequal consequences of mass incarceration for children. *Demography*, 54(1), 361-389. <https://pubmed.ncbi.nlm.nih.gov/28063011/>
- Wood, E. (2022, February 4). A Family Member in Prison? You are Not Alone. Kiplinger.com.
<https://www.kiplinger.com/personal-finance/604154/a-family-member-in-prison-you-are-not-alone>

The Potential Use of Combined High-Frequency Repeated Transcranial Magnetic Stimulation (HF-rTMS) and Virtual Reality (VR) for Post-Stroke Rehabilitation – A Literature Review By Harini Prabhu Shankar

Abstract

VR and rTMS as combined treatment is an inventive approach to stroke rehabilitation that helps in motor and cognitive recovery. The approach has been previously studied, which only focuses on the use of LF-rTMS and VR, with a lack of studies supporting HF-rTMS and VR. HF-rTMS and VR can improve the quality of life factors in stroke patients by dramatically improving motor, cognitive and psychological functioning in comparison to LF-rTMS due to its role in stimulating the damaged hemisphere and increasing its neural activity. Articles/studies on the use of both treatments as standalones as well as the combined use of LF-rTMS and VR were used to hypothesize the potential application of HF-rTMS and VR for stroke rehabilitation. HF-rTMS and VR is still a gray area that requires further research but can potentially improve the lives of stroke survivors/patients by increasing the extent of motor recovery, alleviating psychological/psychiatric symptoms as a result of strokes, shortening the rehabilitation process and establishing long-term neural changes. This paper seeks to overcome these limitations by combining the technologies and investigating its potential uses.

Introduction

Strokes are defined as the acute onset of neurological deficit caused by a deficit of blood flow to the brain. Most commonly caused by blockages, known as an ischemic stroke, the lack of blood results in the sudden death of neurons and decreased oxygen supply to the parts of the brain, as illustrated in *Fig 1*. Strokes may also be hemorrhagic, or caused by bleeding from a blood vessel within the brain. They are one of the leading causes of death and disabilities worldwide (Kuriakose and Xiao). There are many risk factors which lead to strokes, such as hypertension, diabetes, diet, smoking, alcoholism, genetics, age, sex and ethnicity (Nazir). It can result in paralysis, dysphagia, cognitive impairment, speech impairment (otherwise known as aphasia), and various psychiatric illnesses (“Stroke: Hope Through Research”).

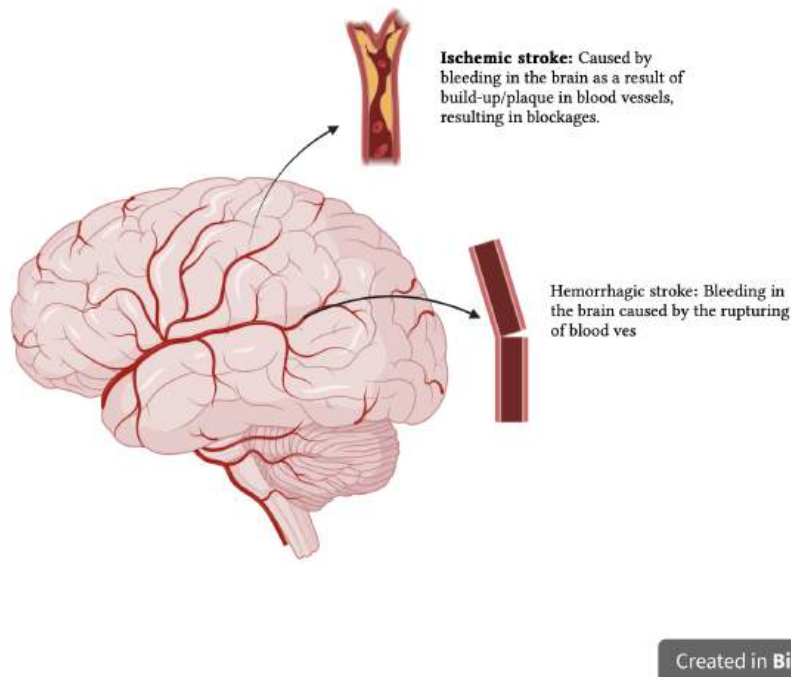


Fig 1: Types of strokes.

Current treatments for ischemic strokes include Alteplase, which have severe side effects such as the potential for life-threatening bleeding. and it is not accessible by all due to its inclusion/exclusion criteria (Nazir). In order to promote improvement in the consequences of strokes, stroke rehabilitation was introduced, a holistic approach to improve the physical, cognitive, neurological and psychological impairments that may impact the quality of life. In the recent decades of research and technological advancements, many new treatments have been proposed and used, proving to be helpful in the journey of recovery from the effects of stroke.

Among these, Virtual Reality (VR) has seen a surge as an optimal rehab treatment. VR is a non-invasive brain stimulation tool (NIBS), providing patients with a simulated 3D environment that is visually immersive. These environments attempt to replicate the real world and provide patients with tasks that challenge their disabilities, therefore improving their impairments (Nazir). A meta-analysis' results have shown that VR is effective in the improvement of both upper and lower-limb motor functioning, balance, gait and daily functioning of post-stroke patients (Zhang et al.).

Repetitive Transcranial Magnetic Stimulation (rTMS) is also a type of NIBS characterized by the repetitive application of a stimuli on a specific cortical region of the brain (Fisicaro et al.). This regulates cortical excitability and therefore, re-shapes the brain plasticity (Sheng et al.). By using an electromagnetic coil placed against the scalp, patients can receive varying levels of frequencies to brain areas, including High-frequency rTMS (HF-rTMS) (>1 Hz) signals, which increases the cortical excitability, and Low-frequency rTMS (LF-rTMS) (≤ 1

Hz) signals which decrease excitability (Fisicaro et al.). The Interhemispheric Inhibition model (IHI) provides insight into what frequency rTMS must be administered to patients, as the model proposes the idea that there is an imbalance between the ipsilesional, the damaged region, and contralesional, the undamaged hemisphere after a stroke, the inhibition from the ipsilesional hemisphere overrides the activity from the contralesional hemisphere, causing motor damage in a specific region of the brain and affecting its related function, as illustrated in *Fig 2*. The general functions of the parts of the brain are, as a result, affected due to the damage caused.

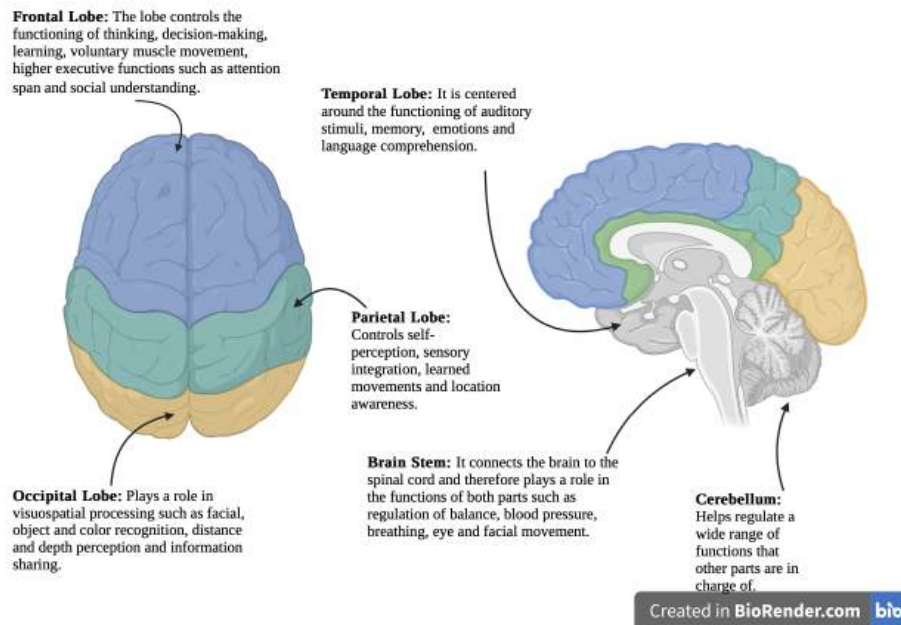


Fig 2: Parts of the brain and their respective functions

Both frequencies can be used to recover motor functionality. LF-rTMS provides an inhibitory effect, reducing the overactivity, while HF-rTMS can provide an excitatory effect, enhancing activity from the damaged hemisphere (Yuan). Both types of rTMS are effective rehabilitation treatments for stroke. *AV et al.* affirms the efficacy of LF-rTMS with significant changes in the excitability of both damaged and undamaged regions, resulting in increased upper limb recovery. Similarly, *Vabalaite* shows a significant improvement in upper limb extremity function by the use of HF-rTMS.

VR and rTMS can be used together in order to further elevate the recovery benefits they provide. VR for stroke rehabilitation requires substantial patient motivation, which may limit its utility. Mental health issues such as post-stroke depression might impede the potential recovery. Hence, rTMS can be used in combination to mitigate the negative psychological symptoms (Sheng et al.). By increasing excitability through rTMS, the brain is more receptive towards the specific task by VR. VR also provides real-time feedback and repetitive tasks that can help reinforce the neural changes caused by rTMS. Several studies have shown the use of LF-rTMS and VR combined for stroke rehab, promoting motor and limb recovery, increased ipsilesional

activation (Johnson et al.; Zheng et al.). However, there are a lack of studies and literature reviews that provide insightful knowledge and evidence on the combined use of HF-rTMS and VR for the increase in quality of life factors in stroke patients.

Therefore, this literature review aims to explore VR and HF-rTMS as standalone treatment methods and hypothesize their potential use and effects as a combined treatment, providing insight into the expansion of stroke rehabilitation treatment methods to improve the quality of life of post-stroke patients.

Virtual Reality (VR) Therapy

VR is broadly divided into 3 categories: non-immersive, where the user views images through a computerized setup; semi-immersive, where the user view images created through “*stereoscopic projections or displays with a fixed visual perspective*” (Demeco) and fully-immersive, where the participant operates within a 3D simulated environment. The virtual environment can also be presented through different methods, such as headsets, large screen projection systems and cave automatic virtual environments (Demeco). Results from VR depend on the type of tasks being provided to the patients. Specialized systems that are tailored to motor and neuroplasticity processes are significantly more efficient as they manipulate practice conditions by representing the patient through an avatar that replicates their movement and therefore, promotes movement in the impaired limb (Nazir). Examples of specialized VR systems are NeuroR, EVA, BTS-Nirvana, Saebo VR, MindMotion Pro and Bi-Manu Trainer, summarized in *Table 1* (Nazir; Bui et al.). Non-specialized VR systems, which include popular console gaming devices such as Nintendo, Wii, Xbox, and Oculus, are also widely used methods for recovery, providing standard movement and motor recovery (Ortiz-Huerta et al.). These systems integrate feedback in the form of scores, rewards, and timings as a means of encouraging patient progress (Bui et al.).

Reference	Participants	Procedure	VR System	Functioning focused on	Results	Limitations
(Kiper et al.)	78 ischemic and 58 hemorrhagic stroke survivors	The experimental group received Reinforced Feedback in Virtual Environment along with conventional rehabilitation. Control group received conventional rehabilitation	A computer connected to a 3D motion tracking system and a high-resolution liquid crystal display projector which displayed the virtual scenarios	Fingers, hands, shoulders, elbows and forearms.	Both groups improved, but significant improvement from the experimental group was seen in F-M UE, FIM, NIHSS, ESAS, time, speed and peak.	It is inconclusive to determine to what extent VR or Occupational therapy contributed to the improvement. The exercises used in the treatment were not controlled, hence it cannot be concluded which parameters were

		alone.	(Non-immersive).			useful.
(Fusco)	23 stroke survivors	Experimental group underwent lower limb training with VR feedback for 2 weeks, and 2 weeks following without VR. Control group performed the same in the opposite order.	Robotic device Omega with VR feedback. (Semi-immersive)	Lower limb	Both groups improved, the experimental group experiencing an early process of functional and cognitive recovery.	Studied a small sample size.
(Mekbib et al.)	23 ischemic or hemorrhagic stroke survivors	The experimental group received VR training and occupational therapy while the control group received occupational therapy alone.	Mirroring neuron VR rehabilitation, HTC Vive tracking stations, leap motion and a central controller were used.	Upper limb extremity	Both groups improved, the experimental group showing increased FM-UE results and neural activity shown by the fMRI.	Small sample size, no follow-up conducted to determine the longevity of results and the control group who underwent occupational therapy did not have an MRI scan and hence, no differential effects on neurological changes were mentioned.
(Proffitt et al.)	1 chronic stroke survivor with left upper extremity hemiparesis.	The participant underwent a 6-week home therapy computer game intervention.	Nintendo Wii remotes controls (Non-immersive).	Upper limb extremity and activities of daily living.	Patient had an increase in quality of life factors, was able to perform a range of daily activities.	It was a case study on 1 individual, hence findings cannot be generalized. The patient was also 17 years post stroke, hence it only studies a specific stage.

Table 1: VR therapy studies for motor recovery.

While there is scientific evidence that justify VR as a successful and functional form of intervention for stroke rehabilitation, there are studies that suggest otherwise, where its results are no different compared to other standard rehabilitation procedures such as occupational therapy (J M de Rooi et al.). VR is highly dependent on optimized device function, and so errors

such as glitches, freezing and inaccurate movement tracking can cause a decrease in motivation, hindering the benefits it may provide (Pallesen et al.). Overall, VR is a popular intervention method for post-stroke rehabilitation and aids in the improvement of motor impairments. However, further research is required in order to expand on its feasibility to make it more effective in introducing neural changes that can improve impaired functioning in stroke patients.

High-frequency Repeated Transcranial Magnetic Stimulation (HF-rTMS) Therapy

Repeated Transcranial Magnetic Stimulation (rTMS) is administered to patients with the goal of stimulating the motor cortex, which plays a role in voluntary regulation of movement (Tomeh et al.; Oroz et al.). The coil is placed on the scalp, controlling the electric currents to stimulate the brain, as shown in *Figure 3*.

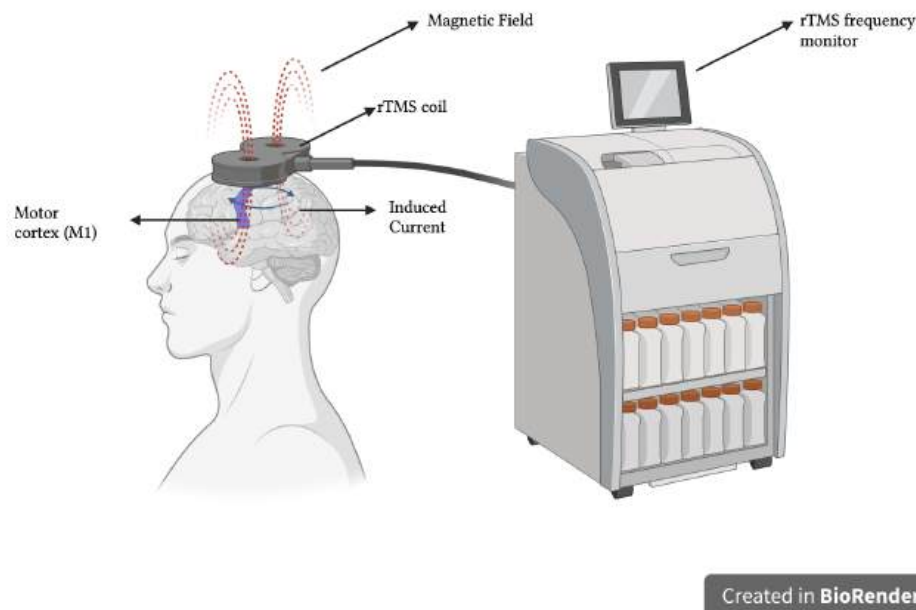


Figure 3: rTMS Procedure. The coil is placed on the scalp of the affected region, administering magnetic stimulation of the required frequency, as displayed on the monitor. This creates a magnetic field perpendicular to the coil, targeting synaptic connections between brain cells in regions that are observed to have decreased or impaired activity.

The intensity of these stimulations is subjective to the patient, determined by motor threshold. The resting motor threshold is a parameter measured to determine the lowest intensity of the motor cortex (M1) stimulation to produce a muscle response to nerve stimulation. It is required for providing the global excitation state of the central core of M1 Neurons. rTMS administered in both high and low frequencies has shown increased repairment of cortical excitability (Fisicaro et al.).

HF-rTMS is widely used for its motor repair and to alleviate the negative symptoms of psychological or psychiatric illnesses. In the field of stroke recovery, however, HF-rTMS is not as widely used in comparison to LF-rTMS due to the likeliness of its adverse effects, such as seizures, headaches, body pain, and hearing problems ((Rossi et al.; Vallejo et al.). In *Li et al.*, findings show out of 153 participants, 6 noted no effect and 4 dropped out due to experiencing pain from the stimulation.

Though HF-rTMS is a less sought-after method, it improves the quality of life for stroke patients. *Mu* studied the distinction between HF-rTMS and LF-rTMS, concluding that HF-rTMS contributes more to the functional connectivity reorganization of the ipsilesional motor network, providing greater benefits for recovery. The increase in cortical excitability and motor-evoked fMRI activation in the ipsilesional motor core results in the restoration of normal ipsilesional M1 activity levels, enhancing motor recovery (Du et al.). These studies have investigated the use of HF-rTMS for stroke rehabilitation as summarized in *Table 2*.

Reference	Participants	Procedure	Functioning focused on	Results	Limitations
(Wang et al.)	45 participants with ischemic or hemorrhagic stroke.	Participants were divided into 3 groups: HF-rTMS, LF-rTMS and a sham group. Each participant received rTMS daily for 2 weeks. HF-rTMS: 10Hz for 1 second per session, 100 sessions each treatment. LF-rTMS followed the same with 1 Hz. The sham group received a fake procedure, with the coil rotating around their scalp producing noises similar to those of actual rTMS.	Motor function	HF-rTMS showed a significant increase in the Fugl-Meyer Motor Assessment score and contralesional hemisphere cortical excitability. No difference in improvement was seen between the LF-rTMS and sham.	Relatively low sample size was used, the electrophysiological indicators used were not sensitive enough for those with hemiplegia. Hence, it might not be effective for patients with severe impairments. It did not focus on the size and location of damage done by the stroke.
(Kim et al.)	133 women, 49 in the HF-rTMS group and 89 in the control group.	All patients were further divided into left (57) and right (76) hemispheric lesions. rTMS sessions were conducted 3 times a week for four consecutive weeks. Stimulation was performed using a biphasic stimulator (MagPro and ALTMS Remed) with 20 Hz per stimulation.	Cognitive ability, speech, memory and activities of daily living.	The HF-rTMS with the left side lesion showed improvement in cognition and depression while the right side showed improvement in cognition, specifically attention,	The number of rTMS varied across patients. The number of participants in each group was uneven, which led to prevalent results being seen in the control group.

				concentration, registration and recall.	
(Sasaki et al.)	21 patients with hemispheric stroke lesions.	The participants were divided into two groups: HF-rTMS group with 11 participants and sham group with 10. Patients received rTMS for 5 consecutive days, in a total of 10 sessions of 10 Hz per stimulation. The sham group received a sham stimulation that replicated sounds of a rTMS procedure.	Leg motor area and lower limbs.	Both groups improved significantly, however the HF-rTMS group scored higher in the Ability for Basic Movement Scale Revised, producing a greater functional recovery.	Small samples of participants were used. No neuro-imaging techniques such as MRI were used to determine the differential effects of the groups. It only studied a specific type of stroke and its severity.
(Cha et al.)	24 patients with subacute stroke.	Participants were divided into two groups, each receiving either HF or LF-rTMS. Each participant went through 20 sessions for 20 minutes, once a day 5 times a week for 4 consecutive weeks. The HF-rTMS groups received 10Hz while the LF-rTMS received 1 Hz.	Motor cortical excitability and balance function.	Both groups showed improvement in the balance function and motor cortical excitability, however, the HF-rTMS group showed increased results in motor cortical excitability. Hence, the study concludes HF-rTMS is effective for cortical excitability.	Relatively small sample size was taken, no neuroimaging techniques such as fMRI was used and a follow-up was not conducted to determine the long-term effects.

Table 2: HF-rTMS therapy studies for motor recovery.

Additionally, HF-rTMS is widely used to treat psychiatric disorders, specifically for those who are resistant towards other treatments, including medication (Wang et al.; Mishra et al.). For disorders like depression, it can target the dorsolateral prefrontal cortex, similar to conventional treatments (Bossuyt). Overall, HF-rTMS has the potential to not only help in motor recovery, but it can also help in alleviating symptoms of psychiatric illnesses that may arise as a result of a stroke, such as post-stroke depression.

The potential combined use of HF-rTMS and VR

The combination of rTMS and VR is an ongoing studied treatment method, wherein it is hypothesized to have complementary benefits on cortical excitability and neural recognition (Chauhan and Das). However, research and evidence supporting their combined use mainly focus on LF-rTMS and VR (Johnson et al.; Zheng et al.). Though research has been done on their combined effects, these studies administer both separately rather than simultaneously, by first administering rTMS, followed by a time gap before proceeding with the VR tasks. The studies concluded the combined use can lead to significant improvements in recovery demonstrated through increased ipsilesional activity and in upper-limb functioning, living activity and quality of life factors (Johnson et al.; Zheng et al.). The comparatively increased number of evidence for LF-rTMS and VR might be due to the fact that LF-rTMS is considerably safer and tolerable in stroke patients, as it has a decreased chance of adverse effects such as seizures, headaches and other bodily pains (Rossi et al.). In combination with VR, which has side effects such as motion sickness, headaches and dizziness (Nazir), HF-rTMS might not be a viable option due to the chances of side effects outweighing their benefits.

Though they both are plausible and effective treatments alone, HF-rTMS and VR both have limitations that can be mitigated by combining the two. For instance, motivation is a key factor in maximizing outcomes from VR-based therapy (Bui et al.). In cases where patients might be diagnosed with post-stroke depression, HF-rTMS, which has been used to alleviate symptoms of psychological illnesses (Bossuyt; Wang et al.), can stimulate the brain and cause neurological changes. These changes can help patients have both the ability and willingness to utilize VR as an intervention and hence, will help enhance their quality of life. As HF-rTMS is used to activate the ipsilesional region, the neural changes that are brought up can be strongly established by utilizing VR to provide repetitive simulated tasks/environments that a patient can go through in order to normalize the movement of the affected limb (Yuan; Nazir). Therefore, it can be said that HF-rTMS and VR combined can help increase and maximize the results of motor recovery as an increase and rebuild of the ipsilesional hemisphere results in restoring the balance between both hemispheres, allowing patients to recover lost functioning at a faster pace compared to the treatments as standalones, as well as LF-rTMS, which targets the overactivity in the undamaged hemisphere. The combination of the two therapies not only result in greater recovery, it also accelerates the rehabilitation process and therefore, allowing patients to produce more robust functional improvement, normalizing their movement and daily activities of living in a shorter period of time. The treatments can also help in targeting specific affected limbs due to the HF-rTMS' ability to target a specific part of the ipsilesional hemisphere as well as VR's specialized systems for specific motor and cognitive functions.

Sánchez et al. explores this combined use of VR and HF-rTMS by utilizing brain-computer interfacing VR along with bilateral rTMS, that is administered on both sides of the brain. Though results are yet to be published, it concludes that its trial will be able to *“provide sufficient evidence supporting VR as an adjuvant therapy along with a known effective neuromodulation approach (rTMS) will enhance the cortico-subcortical activity, therefore*

enhancing motor recovery, reducing the time taken for hand rehabilitation and preventing long-term disability”.

Though, in theory HF-rTMS and VR combined can help improve and accelerate the rate of recovery, the lack of scientific evidence weakens the stance taken. Hence, evidence and findings that explore the topic must be established in order to determine if their combination is effective and determine specific protocols and measures that must be taken in order to deem it safe to be utilized as an official treatment, due to the increased risk of adverse effects with HF-rTMS.

Discussion

This review explores the uses of both HF-rTMS and VR as standalones for post-stroke recovery in aspects such as upper and lower limb recovery, psychological and cognitive improvements. VR provides a realistic and immersive environment that helps stroke survivors to initiate and regulate movement in the affected limbs by allowing them to view the respective limbs through VR in real-time. Additionally, HF-rTMS helps stimulate activity in the ipsilesional hemisphere, resulting in increased motor and cognitive recovery (A.E et al.). The combined use of rTMS and VR is a well-known treatment method, specifically LF-rTMS and VR. Results show that their combined use has proven to significantly improve upper-limb recovery, daily living activity and overall quality of life. However, HF-rTMS and VR and a combined intervention for stroke rehabilitation is insufficiently studied due to its potential side effects that might be brought up such as dizziness, nausea, seizures and other pains. Nevertheless, their combined use may have a positive impact due to their synergy, as they can make up for the latter's limitations and hence, further enhance the recovery process. Due to HF-rTMS role in stimulating the damaged areas affected by the stroke, VR can help establish the surge in excitability exhibited and hence, establish new motor and cognitive changes, therefore improving the rehabilitation process for the long run. HF-rTMS is also widely used for alleviating psychiatric/psychological disorders and hence, can potentially help patients recover from the disorders/illnesses that arise from strokes, such as depression, anxiety and mania. Both treatments combined will not only help improve physical function, but help patients recover from cognitive and psychological deficits and therefore, increase the quality of life factors of these patients. In comparison to LF-rTMS, which targets the overactivity in the contralesional hemisphere, HF-rTMS and VR can speed up the process of recovery due to HF-rTMS role in the ipsilesional hemisphere. Overall, HF-rTMS and VR as a combined protocol for stroke rehabilitation can be effective and even provide stronger improvements for patients in comparison to the therapies as standalones. Though this might be applicable in theory, the lack of sufficient evidence to prove its efficiency results in the potential combination being a tentative treatment that is yet to be applicable (A.E et al.).

Future Directions

The field of stroke rehabilitation and its current treatment method still require further research and exploration in order to simplify the process while enhancing the effects they provide. Factors such as number of sessions for both rTMS and VR, the time period between both treatments, the type and severity of the stroke, time period after the stroke occurred and how these parameters are taken into account when administering these treatment methods still has space for further investigation (Li et al.). Furthermore, VR is limited in its effect on lower-limb recovery and hence, might not be effective in improving its movement. Hence, further research into these listed factors must be done in order to make the rehabilitation process effective and applicable for a wide range of stroke survivors. Moreover, more longitudinal and prospective studies are required to assess the long-term effects of these treatment methods. VR and rTMS can also expand on its viability by more research being done on determining standardized protocols/procedures that target various diverse groups and demographics, in order to make it more accessible to all stroke survivors. These groups can include varying ages, genders, diets and different health concerns/conditions that might impede or be impacted as a result of these interventions. More research is required in the study of the combined use of VR and rTMS, especially with HF-rTMS and VR as currently, there is no reliable evidence that can ascertain that their combined use is effective in improving the quality of life in stroke patients and be considered an effective stroke treatment method. As the primary reason for such a treatment method not being implemented in clinical settings is due to the risk of severe side effects, the procedure of HF-rTMS can be further explored in order to minimize these side effects and make it a safer therapy form (Lieshout et al.). Similarly, VR also harbors side effects such as dizziness and nausea, which are mild but can severely affect stroke patients that have conditions such as epilepsy and seizures (Nazir). The practical application of the combined use of HF-rTMS and VR is limited due to the lack of specialized centers that provide both therapies and standardized procedures to effectively carry out both treatments. Hence, hospital and rehabilitation centers for stroke might find it difficult to successfully implement these treatments for patients to truly benefit from (Nazir).

Overall, further evidence and research is required to expand on the use of HF-rTMS and VR as an intervention to make it applicable and effective for all demographics. The treatment methods, whether individual or combined, have the potential to help stroke patients and survivors improve their lives and expand on the potential of different treatments for stroke rehabilitation.

Limitations

There is no concrete evidence that can support the efficiency and effects of using HF-rTMS and VR combined. While it is hypothesized to have improved results, current evidence makes it inconclusive. Secondly, the studies used in this review are limited by small sample sizes that might not be sufficient in generalizing its results to all stroke patients. In the 8 studies analyzed in *Tables 1 and 2*, 5 were limited due to their small sample size, with an average of 27

participants. The VR studies summarized in *Table 1* lack the use of neuroimaging techniques such as MRI, fMRI, etc. to observe the extent to which VR is effective in recovery and hence, it is difficult to conclude if VR contributes to the rehabilitation process.

This review curated articles from English databases and hence, it might not include studies under the same topic of research that can provide evidence. Furthermore, the review does not take into account the type of stroke and its severity specifically, such as sub-acute stroke for instance, and hence the results and conclusions derived might vary and not be accurate to a specific type.

Conclusion

In conclusion, the combined use of HF-rTMS and VR is a promising field of study that can help expand the variety of treatment methods used in stroke rehabilitation, making the process easier and successful for the patients. Though it is a current area of research that requires more evidence, it has the potential to aid patients in their stroke rehabilitation process and help them restore lost and impaired functioning.

Acknowledgments

Completing this paper would have only been possible with the help from Polygence, giving me this wonderful opportunity to explore and study a topic of my interest. I would also like to thank Dr. Ti’Air Riggins, my mentor who was nothing but supportive throughout the entire process and Mr. Michael Martinez M.D, who had also guided me throughout with thorough feedback

Works Cited

- A.E, Khrulev, et al. “Modern Rehabilitation Technologies of Patients with Motor Disorders at an Early Rehabilitation of Stroke (Review).” *Modern Technologies in Medicine*, vol. 14, no. 6, 2022, pp. 64-78, <https://pubmed.ncbi.nlm.nih.gov/37181286/>.
- AV, Blesneag, et al. “Low-frequency rTMS in patients with subacute ischemic stroke: clinical evaluation of short and long-term outcomes and neurophysiological assessment of cortical excitability.” *National Library of Medicine National Center for Biotechnology Information*, July-September 2015, <https://pubmed.ncbi.nlm.nih.gov/26351545/>. Accessed 4 August 2024.
- “BioRender.” *BioRender: Scientific Image and Illustration Software*, <https://www.biorender.com>. Accessed 1 September 2024.
- Bossuyt, Axel. “HF-rTMS treatment in medication-resistant melancholic depression: results from 18FDG-PET brain imaging.” *PubMed*, August 2009, <https://pubmed.ncbi.nlm.nih.gov/19890238/>. Accessed 19 August 2024.
- Bui, Julie, et al. “Enhancing Upper Limb Rehabilitation of Stroke Patients With Virtual Reality: A Mini Review.” *Frontiers*, 8 November 2021, <https://www.frontiersin.org/journals/virtual-reality/articles/10.3389/frvir.2021.595771/full>. Accessed 10 September 2024.
- Cha, Hyun-Gyu, et al. “Effects of High Frequency Repetitive Transcranial Magnetic Stimulation on Function in Subacute Stroke Patients.” *KoreaScience*, 30 June 2014, <https://koreascience.kr/article/JAKO201418964310124.page>. Accessed 22 August 2024.
- Chauhan, Priya, and Sanjib K. Das. “Efficacy of Simultaneous Application of Repetitive Transcranial Magnetic Stimulation and Virtual Reality Training on Sensory-motor and Cognitive Deficits among Stroke Patients: A Protocol for a Randomised Controlled Trial.” *Journal of Clinical and Diagnostic Research*, 1 July 2024, [https://jcdmr.net/articles/PDF/19686/68577_CE\[Ra1\]_F\(IS\)_QC\(RD_RDW_IS\)_PF1\(RI_DK_OM\)_PFA_NC\(KM\)_PN\(KM\).pdf](https://jcdmr.net/articles/PDF/19686/68577_CE[Ra1]_F(IS)_QC(RD_RDW_IS)_PF1(RI_DK_OM)_PFA_NC(KM)_PN(KM).pdf). Accessed 23 August 2024.
- Demeco, Andrea. “Immersive Virtual Reality in Post-Stroke Rehabilitation: A Systematic Review.” *MDPI*, 3 February 2023, <https://www.mdpi.com/1424-8220/23/3/1712>. Accessed 1 August 2024.
- Du, Juan, et al. “Effects of high- and low-frequency repetitive transcranial magnetic stimulation on motor recovery in early stroke patients: Evidence from a randomized controlled trial with clinical, neurophysiological and functional imaging assessments.” *ScienceDirect*, 3 December 2018, <https://www.sciencedirect.com/science/article/pii/S2213158218303681>. Accessed 19 August 2024.
- Fisicaro, Francesco, et al. “Repetitive transcranial magnetic stimulation in stroke rehabilitation: review of the current evidence and pitfalls.” *NCBI*, 25 September 2019, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6763938/>. Accessed 1 August 2024.

- Fusco, Augusto. "Virtual Reality and Lower Limb Rehabilitation: Effects on Motor and Cognitive Outcome—A Crossover Pilot Study." *MDPI*, 20 April 2022, <https://www.mdpi.com/2077-0383/11/9/2300>. Accessed 18 August 2024.
- J M de Rooi, Ilona, et al. "Effect of Virtual Reality Gait Training on Participation in Survivors of Subacute Stroke: A Randomized Controlled Trial." *NCBI*, 16 February 2021, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8122468/>. Accessed 23 August 2024.
- Johnson, N. N., et al. "Combined rTMS and virtual reality brain-computer interface training for motor recovery after stroke." *PubMed*, 15 February 2018, <https://pubmed.ncbi.nlm.nih.gov/28914232/>. Accessed 17 August 2024.
- Kim, Jongwook, et al. "Effect of Cognition Recovery by Repetitive Transcranial Magnetic Stimulation on Ipsilesional Dorsolateral Prefrontal Cortex in Subacute Stroke Patients." *Frontiers*, 31 January 2022, <https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2022.823108/full>. Accessed 22 August 2024.
- Kiper, Pawel, et al. "Virtual Reality for Upper Limb Rehabilitation in Subacute and Chronic Stroke: A Randomized Controlled Trial." *Archives of Physical Medicine and Rehabilitation*, 15 February 2015, [https://www.archives-pmr.org/article/S0003-9993\(18\)30099-6/abstract](https://www.archives-pmr.org/article/S0003-9993(18)30099-6/abstract). Accessed 17 August 2024.
- Kuriakose, Diji, and Zhicheng Xiao. "Pathophysiology and Treatment of Stroke: Present Status and Future Perspectives." *NCBI*, 15 October 2020, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7589849/>. Accessed 1 August 2024.
- Li, Jiang, et al. "Effects of different frequencies of repetitive transcranial magnetic stimulation on the recovery of upper limb motor dysfunction in patients with subacute cerebral infarction." *National Library of Medicine National Center for Biotechnology Information*, October 2016, <https://pubmed.ncbi.nlm.nih.gov/27904488/>. Accessed 2 September 2024.
- Li, Xiaohong, et al. "Stroke rehabilitation: from diagnosis to therapy." *Frontiers*, 13 August 2024, <https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2024.1402729/full>. Accessed 4 November 2024.
- Lieshout, Eline CC van, et al. "Exploring the experiences of stroke patients treated with transcranial magnetic stimulation for upper limb recovery: a qualitative study." *NCBI*, 6 October 2020, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7541313/>. Accessed 3 September 2024.
- Mekbib, Destaw B., et al. "A novel fully immersive virtual reality environment for upper extremity rehabilitation in patients with stroke." *PubMed*, 14 January 2021, <https://pubmed.ncbi.nlm.nih.gov/33442915/>. Accessed 19 August 2024.
- Mishra, Biswa Ranjan, et al. "Repetitive transcranial magnetic stimulation in psychiatry - PMC." *NCBI*, October 2011, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3271460/>. Accessed 24 August 2024.

- Mu, Qiwen. "Distinction of High- and Low-Frequency Repetitive Transcranial Magnetic Stimulation on the Functional Reorganization of the Motor Network in Stroke Patients." *NCBI*, 20 January 2021, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7840259/>. Accessed 19 August 2024.
- Nazir, Zahra. "Virtual and Augmented Reality in Post-stroke Rehabilitation: A Narrative Review." *NCBI*, 14 April 2023, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10183111/>. Accessed 1 August 2024.
- Oroz, Ricardo, et al. "Transcranial magnetic stimulation therapeutic applications on sleep and insomnia: a review." *Sleep Science Practice*, vol. 5, no. 3, 2021, <https://sleep.biomedcentral.com/articles/10.1186/s41606-020-00057-9#citeas>.
- Ortiz-Huerta, J. H., et al. "[Efficacy of interventions with video games consoles in stroke patients: a systematic review]." *PubMed*, 16 January 2018, <https://pubmed.ncbi.nlm.nih.gov/29323401/>. Accessed 17 August 2024.
- Pallesen, Hanne, et al. "Patients' and Health Professionals' Experiences of Using Virtual Reality Technology for Upper Limb Training after Stroke: A Qualitative Substudy." *PubMed*, 8 February 2018, <https://pubmed.ncbi.nlm.nih.gov/29593910/>. Accessed 19 August 2024.
- Proffitt, Rachel M., et al. "Use of computer games as an intervention for stroke." *National Library of Medicine National Center for Biotechnology Information*, July-August 2011, <https://pubmed.ncbi.nlm.nih.gov/21914607/>. Accessed 19 August 2024.
- Rossi, Simone, et al. "Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research." *NCBI*, 14 October 2009, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3260536/>. Accessed 19 August 2024.
- Sánchez, Francisco José, et al. "Clinical Effects of Immersive Multimodal BCI-VR Training after Bilateral Neuromodulation with rTMS on Upper Limb Motor Recovery after Stroke. A Study Protocol for a Randomized Controlled Trial." *MDPI*, 21 July 2021, <https://www.mdpi.com/1648-9144/57/8/736>. Accessed 2 September 2024.
- Sasaki, Nobuyuki, et al. "High-frequency rTMS on leg motor area in the early phase of stroke." *Springer Link*, 9 August 2019, <https://link.springer.com/article/10.1007/s13760-016-0687-1>. Accessed 22 August 2024.
- Sheng, Rongjun, et al. "Repetitive transcranial magnetic stimulation for stroke rehabilitation: insights into the molecular and cellular mechanisms of neuroinflammation." *NCBI*, 22 May 2023, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10239808/>. Accessed 17 August 2024.
- "Stroke: Hope Through Research." *National Institute of Neurological Disorders and Stroke*, https://www.ninds.nih.gov/sites/default/files/migrate-documents/stroke_hope_through_research_february_2020_508c.pdf. Accessed 1 August 2024.
- Tomeh, Abdulhameed, et al. "Repetitive Transcranial Magnetic Stimulation of the Primary Motor Cortex beyond Motor Rehabilitation: A Review of the Current Evidence." *NCBI*, 10 June 2022, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9221422/>. Accessed 19 August 2024.

- Vabalaite, Birute. "Effects of High-Frequency (HF) Repetitive Transcranial Magnetic Stimulation (rTMS) on Upper Extremity Motor Function in Stroke Patients: A Systematic Review." *MDPI*, 7 November 2021, <https://www.mdpi.com/1648-9144/57/11/1215>. Accessed 1 August 2024.
- Vallejo, Paula, et al. "Repetitive Transcranial Magnetic Stimulation in Stroke: A Literature Review of the Current Role and Controversies of Neurorehabilitation Through Electromagnetic Pulses." *Cureus*, 11 July 2023, <https://www.cureus.com/articles/155677-repetitive-transcranial-magnetic-stimulation-in-stroke-a-literature-review-of-the-current-role-and-controversies-of-neurorehabilitation-through-electromagnetic-pulses#!/>. Accessed 19 August 2024.
- Wang, Qi, et al. "Effects of high-frequency repetitive transcranial magnetic stimulation over the contralesional motor cortex on motor recovery in severe hemiplegic stroke: A randomized clinical trial." *ScienceDirect*, 2 April 2020, <https://www.sciencedirect.com/science/article/pii/S1935861X20300735>. Accessed 20 August 2024.
- Wang, Qi, et al. "Intensive rTMS for treatment-resistant depression patients with suicidal ideation: An open-label study." *ScienceDirect*, 16 June 2022, <https://www.sciencedirect.com/science/article/abs/pii/S1876201822001873>. Accessed 23 August 2024.
- Yuan, Xiaoxia. "Promotion of Poststroke Motor-Function Recovery with Repetitive Transcranial Magnetic Stimulation by Regulating the Interhemispheric Imbalance." *MDPI*, 2020, <https://www.mdpi.com/2076-3425/10/9/648>. Accessed 1 August 2024.
- Zhang, Bohan, et al. "Virtual reality for limb motor function, balance, gait, cognition and daily function of stroke patients: A systematic review and meta-analysis." *PubMed*, 6 March 2021, <https://pubmed.ncbi.nlm.nih.gov/33675076/>. Accessed 1 August 2024.
- Zheng, Chan-juan, et al. "Effect of combined low-frequency repetitive transcranial magnetic stimulation and virtual reality training on upper limb function in subacute stroke: a double-blind randomized controlled trial." *Springer Link*, 16 April 2015, <https://link.springer.com/article/10.1007/s11596-015-1419-0>. Accessed 4 August 2024.

Using Algae as A Catalyst for Carbon Neutrality By Amber Dogra

Abstract

Over the past 200 years, global carbon concentration has increased by 50%. Carbon Capture and Storage (CCS) is the most recent form of carbon capture technology available to minimize the effects of carbon emissions. Although it is efficient, using carbon capture brings with it many environmental and economic concerns. This is where algae comes in. Algae is much more efficient at absorbing carbon dioxide than other plants during photosynthesis, thus holding significant potential to act as a catalyst for future carbon neutrality. Algae can play a crucial role in alleviating current and future greenhouse gasses using this natural process. Algae can virtually grow anywhere that has sunlight, CO₂, and water. Furthermore, the biomass of algae can be used as a biofuel alternative, in turn allowing for a sustainable alternative to fossil fuels. Integrating algae-based systems into carbon capture and other utilization strategies shows a promising path toward environmental sustainability and carbon neutrality.

Introduction

The idea of carbon capture has greatly evolved since the discovery of the concept of gasses and their properties in the 19th century to the 21st century, where there have been many technological developments that eventually led to the use of carbon capture systems. In the early 1900s, around the time of the Industrial Revolution, many factories used fossil fuels as a means to fuel their production, causing significant amounts of CO₂ to be produced. Scientists started to observe the effects of carbon dioxide in the atmosphere with increased problems of global warming and climate change (Lacis et al 2010) beginning in the 1950s.

In the 1970s and 80s, there was increasing awareness of climate change and greenhouse gas emissions. The idea of capturing CO₂ from power plants and industrial sources gained attention. A decade later, the idea of carbon capture and storage began being discussed in scientific communities as a potential strategy to mitigate climate change. This included pre-combustion, post-combustion, and oxy fuel technologies. The 21st century brought up projects that were launched worldwide to test different carbon capture and storage strategies such as The Tomakomai Project in Japan.

Today, the global atmospheric concentration of carbon dioxide (CO₂) is ~420 ppm (Bai et al 2024). The majority of the recent increase in CO₂ is contributed by human activities like the burning of fossil fuels and the reduction of forest cover (deforestation) (IEA 2023). In fact, as of 2023, over 80% of energy used worldwide was sourced from burning fossil fuels (Perkins, Edwardes-Evans, 2023).

At the moment, the cheapest sources of energy are solar and wind energy; these are clean energies (energy produced by resources that will not run out and the byproducts have minimal to no effects on the environment.). Since the 1900s, there was a quick switch from wood and charcoal to coal (A non-renewable fossil fuel composed of primarily ancient plant matter, used for energy, electricity, or industrial processes, it emits high levels of CO₂), for industrial

purposes, especially in factories, ships, and trains. According to Jean-Paul Rodrigue (et al) Geography of Transport Systems (2006), the energy density of wood is around 17.1 MJ/kg (megajoules per kilogram) and coal is 31.4 MJ/kg, making it almost twice as energy-dense. The fossil fuel supply is projected to be depleted by 2112, with coal being the only form of fossil fuel after the year 2042 (Shafiee, Topal 2009).

Carbon dioxide is a highly potent greenhouse gas, meaning that once it's in the atmosphere, it traps heat and prevents it from getting out. According to the American Geoscience Institute, Coal releases the most amount of CO₂ at 228.6 pounds of CO₂ emitted per million British thermal units, compared to natural gas at 117 pounds of CO₂ emitted per million British thermal units (AGI, 2024). This contributes greatly to climate change: polar caps are melting, natural disasters are more extreme, rising sea levels, and the imbalance of the population of organisms in the environment. It would be difficult for humans to stop using fossil fuels due to their deep integration into our daily lives. Things like transportation, cooking, and energy production are primarily achieved by fossil fuel usage. Considering that it would be difficult to stop using fossil fuels for general use in society, humans have considered Carbon Dioxide Removal (CDR).

There are two main methods of carbon dioxide sequestration: Carbon Capture and Storage (CCS), and Carbon Dioxide Removal (CDR) (Subraveti 2023). CDR is the act of catching ambient carbon and storing it for later industrial use, while CCS catches ambient carbon and permanently stores it underground. CCS is the most recent form of carbon capture technology developed to minimize the effects of climate change. Although it is efficient, the carbon capture process is not cost-friendly (Yasemi et al 2023). In contrast, plants are great at converting CO₂ into oxygen, especially algae (Moreira 2016). Since the 1950s, algae has been 10-50 times more efficient at absorbing carbon dioxide than other plants during photosynthesis, thus holding significant potential to act as a catalyst for future carbon neutrality (Batista et al., 2015). Algae can be easily cultivated in wastewater as well as salinated environments, which depletes competition for arable lands.

The biomass of algae can also be used as a biofuel alternative, acting as a 'green diesel' to fuel engines in trucks, heavy machinery, and factories, providing a sustainable alternative to fossil fuels.

The purpose of this paper is to explore and review published literature regarding algae as a potential method to combat climate change, primarily by providing an overview of CCS and algae's potential as a biofuel.

Methods

Key terms used: biofuel, algae, carbon capture, carbon storage, carbon neutrality, photosynthesis, energy, coal, fossil fuels. Only documents published in English were included in this review. There were no restrictions on study type, report type, epidemiological design, or country of origin. Additionally, there was a restriction on publication date, only literature published after 1978 (when algae was found to be a good fuel alternative) was considered for

inclusion. However, the most recent research and data were ascertained for the most accurate intel when applicable. Resources utilized for this paper were found via the Google Scholar Search Engine.

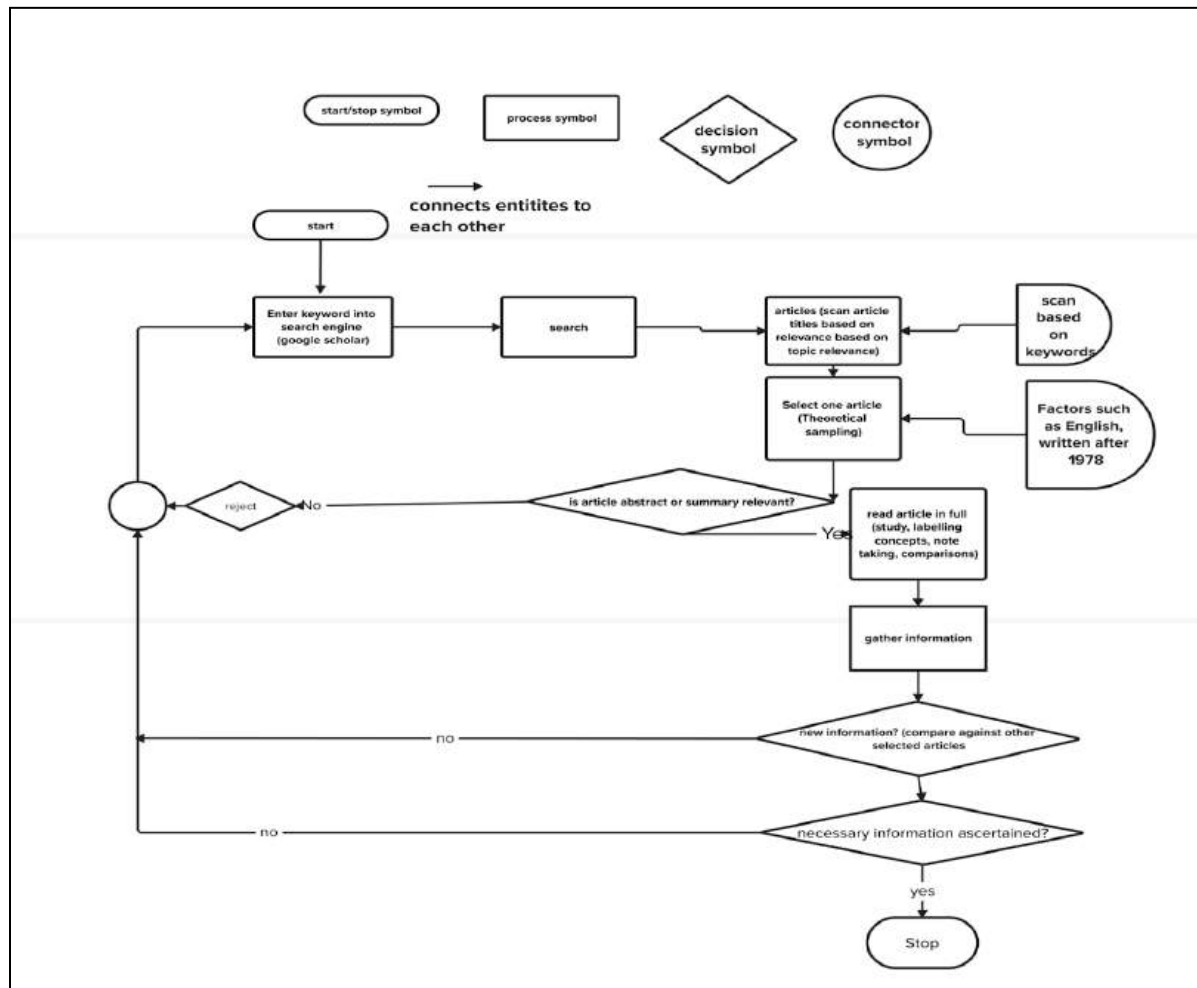


Figure 1. Inclusion and Exclusion Criteria Flowchart

Carbon Capture & Storage

CCS Sequestration is an imperative technology in the fight against climate change. With the amount of fossil fuels being burned by factories, infrastructure, energy, and transportation, we need to find a viable solution to slow down the effects of climate change.

Once anthropogenic greenhouse gasses such as CO₂ are in the atmosphere, they begin depleting the ozone molecules in the stratosphere. Because of this, more UV rays are entering the atmosphere, which may contribute to many health issues (Udoh 2014) Along with this; the density of CO₂ molecules is beginning to trap more and more heat from the sun, contributing to global warming(Elias 2018).

CCS Sequestration is an imperative technology in the fight against climate change. With the amount of fossil fuels being burned by factories, infrastructure, energy, and transportation,

we need to find a viable solution to slow down the effects of climate change. There have been efforts to fight this crisis, such as the Tomakomai Project in Hokkaido Island, a full-scale CCS operation by Japan CCS Co. Ltd (JCCS) beginning in 2008.

The Tomakomai Project was the first full-chain carbon capture and storage project in Japan, which is located right next to an oil refinery to avoid the immediate spread of CO₂. The process begins with the capture of CO₂ using a chemical absorption process from the flu gas (the gas produced by burning fossil fuels, coal, oil, or natural fuels for large-scale power plants) produced by the refinery's hydrogen production unit. After it has been captured, CO₂ is compressed to a higher pressure for it to be transported. It is transported via pipelines to their storage sites. The compressed CO₂ is then injected deep underground into geological formations. The Tomakomai Project used two storage locations: one onshore and one offshore. The one offshore is located beneath the seabed in a sandstone formation. There is a highly stringent monitoring system (the use of many sensors) to track the underground CO₂ and ensure it hasn't leaked

Cons of Carbon Capture

Although the concept of CC is seen as a promising way to mitigate climate change, there are many drawbacks. Factors like economic viability and land usage are problems that should be taken into account when using CC systems.

Closely related to the high costs is the energy required to facilitate this technology. Separating and compressing CO₂ requires significant amounts of energy, which lowers the overall efficiency and output of the powerplant by 10-40% (Cuéllar-Franca & Azapagic, 2015). This energy usage offsets the benefits of emission reductions. Reduced efficiency and output capacity of power plants with CC technology may need more power plants to meet the energy demands, which increases costs, land-use impacts, and other environmental accommodations.

Carbon sediment from CC systems is primarily stored underground in geographical formations or used in industrial processes. However, the capacity of storage sites is limited and there are risks of CO₂ leaking over periods into groundwater, potentially tainting water supplies or causing vegetation damage(Zwaan et al 2009). This tends to be a problem since the point of storing carbon sediment underground is so that it remains there permanently.

History of Algae as a Carbon Capture Method

The earliest work for exploring the use of algae as a method of CC was from the 1950s and 60s by scientists like H.C Gaffron at the University of Chicago. Their research examined the mechanisms of microalgae fixing CO₂ during photosynthesis and the potential to use CO₂ from industrial flue gasses as a source of inorganic carbon. The research highlighted the ability of certain microalgae, such as *Oscillatoria*, to efficiently sequester CO₂ into oxygen.

In the 1990s-2000s concerns were rising over anthropogenic CO₂ emissions, and the associated climate impacts sparked a resurgence in interest in microalgae to be used as a biological carbon capture approach. There were many major initiatives involved, such as Japan's

artificial upwelling and cultivation system tested in the 1990s, the U.S. Department of Energy's research on microalgae CO₂ sequestration in the late 1990s, and the formation of companies like Greenhouse Gas Solutions (GGS) in the late 2000s commercializing algal bioprocesses for carbon capture. During this period, many studies evaluated different cultivation methods (open ponds, arable, non arable land), algal species, CO₂ transfer technology, and the overall technical feasibility and CO₂ sequestration potential of using algae with industrial waste gas.

Recently, researchers have been developing specialized algal strains, trying to optimize growth conditions, experimenting with combined approaches for using algae as a CO₂ sequestration method, and integrating algae cultivation near power plants for CO₂ sequestration (Devi et al 2018).

Algae is highly versatile beyond the factor of biomass production since it has a high potential for biofuel production. There are currently global efforts aimed at modifying the intensity of the accumulation of lipids, alcohols, hydrocarbons, and other energy storage methods in photosynthetic organisms, yeasts, and bacteria through genetic engineering (Radakovits et al 2010). Unlike traditional crops, algae require minimal land and other resources, which makes them an attractive option for biofuel production. Microalgae is known for its prolific biomass production and storage capabilities (Majid et al 2014); it thus shows great potential as a frontrunner in the realm of biofuel production. With increasing petrol prices, concerns over the control of oil (Morris et al 2016), and environmental problems such as global warming, microalgae can offer a beacon of hope in the arena of renewable biofuels.

Microalgae thrive in diverse environments. This can be a freshwater pond, coastal region, and even the most polluted waters, so it provides flexibility in cultivation locations. Algae's adaptability mitigates the pressure on arable land and opens up the opportunity to utilize non-arable lands and wastewater for biofuel production (Savage, Hestekin 2013). Additionally, this adaptability to saline or brackish water expands the potential for large-scale production in the face of limited freshwater.

Advancements in biotechnology and genetic engineering show prospects of increasing the productivity of biofuel production. For example, a company named Algenol. This organization works to produce a biofuel that has been made of genetically modified cyanobacteria to convert CO₂ into ethanol through photosynthesis. Researchers are exploring ways to manipulate the metabolic systems of algae to increase its lipid accumulation (Radkovitz et al 2010). Using fixed experiments, scientists aim to optimize different strains of algae to different specific biofuel applications, this may include biodiesel or bioethanol production.

History of Algae as a Biofuel

Algae was first discovered as a fuel alternative in 1942 by Harder and Von Witsch. Later on in 1974, under the leadership of President Jimmy Carter, researchers studied the simple growth requirements of algae to commercially produce it for fuel during the oil crisis. A study showed that if microalgae biomass from open ponds were to be produced at the level of 23 grams

per square meter on 3.18 billion m^2 of the land area of the southwestern desert land, it could supply 100% of the current United States demand for transport oil (Sheehan et al 1998).

Algae as a Biofuel

With the mass production of algae as a Carbon Capture method, it is important to acknowledge the biomass that would be produced by it. This biomass could be integrated into biodiesel: a cleaner biofuel that is derived from plants. The process is transesterification, which involves taking the triglyceride oil which makes up most oils and fats, then processing it with alcohol and a catalyst which will separate the glycerides (Radkowitz et al 2010). The result is a clean-burning, simple alkyl esters biodiesel and a usable glycerol byproduct. With the ever-growing population and increasing demand for fuel, this industry has a very high potential to quickly grow and surpass the cost and performance of petroleum diesel.

In 2005, the US government announced a mandate to switch over to ultra-low sulfur diesel, which is ~15 ppm (parts per million). This was made to be accompanied by reducing soot/Particulate Matter (PM) by 90% and nitrogen oxides (NO_x) by 96% by 2010 (EPA 2006). Algae has the potential to exceed these standards, but until that can be proven, a renewable, clean-burning fuel such as biodiesel (an alternative fuel based on vegetable oils or animal fats) is the first target. It is a cleaner-burning replacement for petroleum diesel fuel and has shown a significant reduction of soot, particulate matter, and CO.

Diesel fuel is a product used in many industries because of its low explosive, ignition-powered motor fuel, which is why the primary focus is to make algae diesel. Unfortunately, the existing biomass-derived diesel has some problems. It has the same NO_x emissions as vegetable oils and produces 10% more CO₂ emissions than gas. As for the sulfur content, it is the same as that in gasoline, which defeats the purpose of finding an alternative to reduce the US dependency on gasoline. The advantage of algae is that there can be a specifically tailored oil type for whatever the need may be. With the right resources, algae can create fuel to meet the US environmental regulations and the needed performance for a cheaper cost.

Conclusion

Escalating global carbon emissions, which are surpassing 400 ppm worldwide, overshadows the fact that there is an urgent need for sustainable and efficient mitigation strategies. While Carbon capture technology is available to sequester this carbon, there are many environmental and economic concerns surrounding the usage of Carbon capture technology. Algae may emerge as a potential agent in the quest for carbon neutrality.

Algae are highly efficient at sequestering carbon dioxide compared to other plants; along with this, it is easily cultivated and can be mass-produced. For the biomass, it may be used to produce a biofuel. This makes it a multifaceted solution to the economic and environmental concerns surrounding carbon capture technology.

The integration of algae-based carbon capture and the utilization of algae for biofuel has the potential to offer a more holistic approach to environmental sustainability. Further research must be done to overcome some challenges connected with large-scale cultivation, including optimizing growth conditions, improving harvest techniques, and enhancing the efficiency of biofuel production. As technology advances and our understanding of the algal system deepens, the role of algae in reaching carbon neutrality may become increasingly significant.

Acknowledgments

This paper was written with the intent of receiving college credits from the University of California Irvine under the supervision of my research advisor John-Paul Angulo, MPH, and the help of the Polygence editing team.

Works Cited

- Savage, P. E., & Hestekin, J. A. (2013). A perspective on algae, the environment, and energy. *Environmental Progress & Sustainable Energy*, 32(4), 877-883.
- Moreira, D., & Pires, J. C. (2016). Atmospheric CO₂ capture by algae: negative carbon dioxide emission path. *Bioresource technology*, 215, 371-379.
- Yasemi, S., Khalili, Y., Sanati, A., & Bagheri, M. (2023). Carbon capture and storage: Application in the oil and gas industry. *Sustainability*, 15(19), 14486.
- Rosa M. Cuéllar-Franca, Adisa Azapagic, Carbon capture, storage and utilisation technologies: A critical analysis and comparison of their life cycle environmental impacts, *Journal of CO₂ Utilization*, Volume 9, 2015, Pages 82-102, ISSN 2212-9820,
- van der Zwaan, B., & Smekens, K. (2009). CO₂ capture and storage with leakage in an energy-climate model. *Environmental Modeling & Assessment*, 14, 135-148.
- Ana Paula Batista, Lucas Ambrosano, Sofia Graça, Catarina Sousa, Paula A.S.S. Marques, Belina Ribeiro, Elberis P. Botrel, Pedro Castro Neto, Luisa Gouveia, Combining urban wastewater treatment with biohydrogen production – An integrated microalgae-based approach, *Bioresource Technology*, Volume 184, 2015, Pages 230-235, ISSN 0960-8524,
- Morriss, A. P., & Meiners, R. E. (2016). Competition in global oil markets: a meta-analysis and review. URL: http://freebeacon.com/wp-content/uploads/2013/04/SAFE-Study_GlobalOilMarkets.pdf (Accessed: 25.10.2014).
- Bai, J., Li, R., Jiang, Y., Zhang, J., Li, D., Cai, Z., & Zhang, Z. (2024). Efficient agricultural water research under elevated global carbon dioxide concentration–Based on bibliometric analysis. *Agricultural Water Management*, 299, 108874.
- Perkins R. & Edwardes- Evans H. (2023) Fossil fuels 'stubbornly' dominating global energy despite surge in renewables: Energy Institute
- Elias, S. A. (2018). Greatly increased CO₂. *Encyclopedia of the anthropocene*, 115-128.
- Radakovits, R., Jinkerson, R. E., Darzins, A., & Posewitz, M. C. (2010). Genetic engineering of algae for enhanced biofuel production. *Eukaryotic cell*, 9(4), 486-501.
- AGI (2024) How much carbon dioxide is produced when different fuels are burned?
- Subraveti, S. G., Rodríguez Angel, E., Ramírez, A., & Roussanaly, S. (2023). Is carbon capture and storage (CCS) really so expensive? An analysis of cascading costs and CO₂ emissions reduction of industrial CCS implementation on the construction of a bridge. *Environmental Science & Technology*, 57(6), 2595-2601.
- Sheehan J., Dunahay T., Benemann J., & Roessler P. (1998) Look back at the U. S. Department of Energy's Aquatic Species Program: Biodiesel from Algae; Close-Out Report
- Onyeaka, H., Miri, T., Oibileke, K., Hart, A., Anumudu, C., & Al-Sharify, Z. T. (2021). Minimizing carbon footprint via microalgae as a biological capture. *Carbon Capture Science & Technology*, 1, 100007.
- EPA (2006) Nitrogen Oxides (NO_x) Control Regulations

- Andrew A. Lacis et al. ,Atmospheric CO₂: Principal Control Knob Governing Earth's Temperature.Science330,356-359(2010).DOI:10.1126/science.1190653
- Majid, M., Shafqat, S., Inam, H., Hashmi, U., & Kazi, A. G. (2014). Production of algal biomass. Biomass and bioenergy: processing and properties, 207-224.
- Udoh, A. O. (2014). Ozone layer depletion: causes and effects. Journal of Research in Pure and Applied Sciences, 3(1), 150-156.
- Devi, A., Saran, C., Ferreira, L.F.R., Mulla, S.I., Bharagava, R.N. (2024). Sustainable Approaches to Algal Biofuels: Opportunities, Key Challenges and Current Status. In: Arya, S.K., Khatri, M., Singh, G. (eds) Value Added Products From Bioalgae Based Biorefineries: Opportunities and Challenges. Springer, Singapore.
https://doi.org/10.1007/978-981-97-1662-3_8
- IEA- World Energy Outlook (2023)
- Rodrigue, et al (2006) The Geography of Transport Systems
- Shafiee, S., & Topal, E. (2009). When will fossil fuel reserves be diminished?. Energy policy, 37(1), 181-189.

The Relationship Between Green Infrastructure and Carbon Emissions: A Literature Review

Review By Srinath R. Dadiredy (student)

Abstract

Climate change is defined as a long-term shift in temperatures and weather patterns, and is mainly caused by the increased production of CO₂ emissions and other pollutants leading to an increase in the greenhouse effect and global warming. With this incredibly pressing and rapidly growing issue, it is important to analyze possible solutions to the problem, a major one being green infrastructure. This paper aims to answer in what ways and how well green infrastructure is able to stop and combat CO₂ emissions through a literature review and analysis of information from other published research papers and experiments. It was found that green infrastructure proved to be very effective at combating CO₂ emissions, and - more importantly - was able to do so using a myriad of techniques and methods. This paper serves to summarize the data found, presenting an original analysis of the data and providing a new perspective on the topic at hand to answer the research question.

Keywords

Engineering Mechanics; Civil Engineering; Green Infrastructure; Carbon Emissions; Air Quality

Introduction

With cities and urban living becoming a more and more popular and mainstream type of living - as much as 75% of the world is predicted to be living in cities by 2050 - it has become increasingly more important to consider the ramifications of such drastic urban growth.⁷ Cities, with their numerous societal benefits such as job opportunities and economic growth, lead many to stand supportive of them and avidly encourage their growth with little thought to much else. However, with the benefits and growth that comes from cities, there also come the downsides and negatives of rapid urban expansion, one of the most major being its effect on the environment. One piece of urbanization that heavily affects the environment are the heightened emissions of carbon dioxide, which have many negative effects such as decreased air quality, increased pollution, as well as speeding the greenhouse effect and global warming.

CO₂ emissions, defined as being emissions of the gas carbon dioxide or other greenhouse gasses through the consumption of any type of matter, are most prevalent in cities. Cities and urban living are behind more than 70% of global CO₂ emissions through their heavy use of emission sources such as electricity and automotive transportation, making them a forefront for the issue of climate change pollution.⁷ These issues have created the need for a solution to help save the environment, a solution to combat emissions; of those that have emerged, one of the more practical has come in the form of green infrastructure. Green infrastructure comes in a variety of sources, but their main purpose is to combat the negative effects or sources of CO₂ emissions to aid the environment. Although many laud green infrastructure as an answer to the

issue of CO₂ emissions, there have been arguments surrounding the viability of green infrastructure from an implementation and economic perspective.

The goal of this research was to determine the relationship between green infrastructure and emissions by analyzing papers from other researchers and summarizing them in a literature review. The review contains information on green infrastructure and emissions, as well as an analysis of the relationship between them. Summarizing the findings, it is clear that a relationship does exist as evidence shows green infrastructure is very effective at offsetting emissions in cities due to its various forms and functions detailed throughout the paper. Through literature review, this research paper will examine the effectiveness of green infrastructure in combating CO₂ emissions by answering the research question: In what ways does green infrastructure effectively offset CO₂ emissions in cities?

What Are Common Forms And Purposes Of Green Infrastructure Found In Cities?

When considering green infrastructure as a possible solution to CO₂ emissions, it is important to consider the different types of green infrastructure and their effect on emissions. Green Infrastructure can come in many forms, including green spaces, parks, rain gardens, urban forestry and agriculture, as well as green roofs and walls.⁵ These various forms all perform various tasks, but they can fall into one of, or a combination of, three categories: reduction, extension, or protection.³ Reduction is meant to reduce emissions, either by removing CO₂ out of the air or preventing their spread at the source. Extension furthers the distance that CO₂ emissions travel, diluting them further in the air. Finally, protection works to remove people from the pollution directly and prevent contact.³ To summarize, reduction removes emissions, extension dilutes them, and protection prevents them from having harmful effects.

There are several steps that green infrastructure can take to meet these goals of offsetting emissions, and they typically come in differing styles and methods. A majority of these methods attempt to mitigate emissions by diluting them in the air to reduce their effect on the environment and improve air quality. Methods vary, but they include using natural barriers such as hedges and trees to break apart winds and dilute the air, using green roofs and walls to take CO₂ emissions out of the air, and separating people from CO₂ sources such as vehicles through natural barriers like bushes or hedges.³ By breaking apart CO₂ emissions and spreading them finer throughout the air, green infrastructure aims to dampen the effects of CO₂ emissions while also making it easier for greenery to absorb and remove it from the environment.² Another major method that green infrastructure uses to battle climate change is the use of water in their designs. Water is often used in green infrastructure to reduce the effects of CO₂ emissions. This can come in the form of heating for homes, as underground water systems have been shown to be capable of heating homes through the reuse of heat from wastewater in sewage tunnels. It has been found that a system like this can reduce carbon emissions from heating by as much as 60%, as well as satisfying heat demands with no emissions for over 60% of the year in countries like the United Kingdom.¹² Furthermore, Sustainable Urban Drainage Systems, or SUDS, are often used to control surges of water, allowing for environmental protection from symptoms of climate

change, such as increasing tropical storms or rising sea levels.¹ In summary, green infrastructure in cities typically use natural barriers and vegetation to stop emissions through absorption, dilution, or separation, while other methods, like water, can be used to a lesser extent to help green infrastructure reduce the effects of emissions, such as the heat-island effect where urban areas experience hotter temperatures than surrounding areas.

What Are The Main Causes And Effects Of CO₂ Emissions In Cities?

Cities house some of the largest and most common producers of CO₂ due to their increased energy consumption and population size. Urban energy usage typically falls into three major categories, being electricity, ground and air transportation. Examples of these include personal or public transit, as well as heating and industrial combustion to produce energy or manufacture products.⁸ This is important as CO₂ emissions tend to scale with population size and energy consumption, two categories which cities place very high in. This makes them a hotspot for emissions, as well as a focus when it comes to stopping them.⁷ Therefore, it is important to note that cities will compound the effects of emissions by scaling up the amount of producers. Cities have higher populations and higher needs, leading to more production and emissions and a consequence.

These emissions tend to be a problem as they negatively affect the environment, seen in cities with their heavy pollution through the form of smog or gas. For example, burning fossil fuels such as gasoline in engines or coal in electrical plants, as well as industrialization such as factories which produce emissions as a byproduct of production can lead to increased smog and pollution. This smog spreads air pollutants such as dust and gas, reducing the air quality which increases health risks in city-dwellers, harms plant life in the environment, and reduces the overall quality of life in cities.⁴ Evidence from earlier states that the emissions that cause the issues scale with cities, meaning the negative effects such as smog and pollution scale as well. This is not to mention the overall effect that CO₂ emissions have on the globe, such as extreme climate change, a decline in biodiversity as a result, and global warming. With more and more people moving into cities, it is important to recognize and combat the issue of CO₂ emissions before it gets out of hand.

Results and Discussion

Reduction

When looking at the effectiveness of green infrastructure against emissions, it is important to realize the differing factors behind that effectiveness. This goes beyond just the function of green infrastructure, going to the base of vegetation they are built upon, as well as nature and effects of emissions themselves. The purpose behind green infrastructure when related to emissions can be divided into three categories: reduction, extension, and protection.³ Using these categories as metrics for success, the effectiveness of green infrastructure can be measured by establishing how well these methods accomplished their respective goals. Research has

shown an inverse relationship between green infrastructure in cities and the emissions seen in those cities in regards to the three aforementioned categories, as CO₂ pollution drastically decreases as more green infrastructure is introduced into the urban landscape. This relationship can most clearly be seen in the reduction aspect: for example, analysis of 55 different cities showed that, in total, the urban tree canopies established were able to remove 711,000 tons of CO₂ pollution annually.⁹ Further research showed that there was a consistent decrease in the average exposure of individuals to pollution in outdoor areas - about 46% - from the implementation of roadside trees in cities.⁹ The trends demonstrate how greenery and implementation of different vegetation is directly related to the reduction of pollution from emissions in cities. Furthermore, looking at the data taken, it can also be seen that proximity of the green infrastructure to the source of the emissions also has a profound effect on its ability to combat emissions as it allows for a more direct route for the absorption of the emissions by the vegetation.⁹ As shown previously, roadside trees were able to drastically decrease the emissions in cities, a feat only possible as the green infrastructure was in very close proximity to the source of emissions, being the vehicles on the road.

Factors Behind Carbon Emissions

It is also important to look at the carbon emissions themselves when gauging the effectiveness of green infrastructure, as the true effectiveness can only be seen when looking at how well green infrastructure combats the worst of the CO₂ emissions. CO₂ emissions can come in many different forms, but there are typically three categories for them in cities, similar to green infrastructure: these are electricity, forms of transportation, and heating through combustion.⁶ In looking at these categories, more metrics for measurements are given to gauge the effectiveness of green infrastructure, as it can be analyzed as how well the infrastructure targets these various types of emissions. Previous examples such as the success of roadside trees and other forms of greenery placed near transportation in lowering emissions once again shows how green infrastructure is able to successfully offset emissions in cities, at least to the extent of transportation. However, digging deeper, it is clear to see the effect does not only stop there. Electricity and heat function in similar ways as electricity is produced through methods that create CO₂ emissions as a byproduct, while heat production can generate CO₂ byproducts, but heat tends to be a result of the CO₂ emissions in general. CO₂ tends to trap heat by preventing air circulation and absorbing heat, leading to densely-populated areas with high emissions, such as cities, to become much hotter than the surrounding area; this same effect is a factor responsible for the greenhouse effect and global warming as a whole, thus showing the importance in combating it. On this front as well, green infrastructure shows success in combating these effects of CO₂. In areas where green infrastructure had been implemented, temperature reductions of up to 1 degrees Celsius were found, with reductions increasing based on the scale of the implementation.¹⁰ Vegetation such as roof canopies or urban forests provide shade in cities, cooling surrounding areas, while generally absorbing heat and removing it from the environment. Further layered by removing CO₂ emissions from the environment, green infrastructure has

proven to be very effective in reducing and mitigating the major effects and sources of emissions, such as the rising temperatures and emission rates in cities.

Extension

The second purpose of green infrastructure, being extension, is the dilution of CO₂ emissions in the air to reduce the potency and effects, which can be used as a metric of measurement for green infrastructure. As a whole, air quality tends to be a large problem in cities due to the volume of CO₂ emissions, which tend to gather in large pockets due to poor air circulation and ruin the air, leading to negative effects on both city-dwellers and the environment.⁴ Extension is a measure to solve that problem, as it aims to make CO₂ emissions travel farther in the air, diluting it and making it easier for green infrastructure to absorb and break down while lessening the effects on both the environment and people. It has been proven once again that green infrastructure is able to accomplish this task through various methods. Green roofs, a type of greenery where different vegetation is planted on top of buildings to create a sort of rooftop forest, have shown to be particularly effective against CO₂ emissions, with a direct relationship having been found between green roofs and emissions. A single green roof leads to a direct air quality improvement of 0.001 kg per square meter annually, a figure demonstrating green infrastructure's effectiveness.⁴ Being far in the sky could give green roofs an edge as they have more access to the emissions in higher airspaces. These figures are possible as green infrastructure is typically used to break apart emissions that travel with air currents: for example, a tree with its many branches and varying paths of airflow serve as a good natural barrier for CO₂ emissions as they break apart winds and emissions into smaller parts that are easier to deal with.³ The use of natural barriers through green infrastructure leads to the effective offsetting of CO₂ emissions as emissions are diluted and broken apart in the air, reducing the negative effects while making it easier for vegetation to further break them down and absorb the emissions.

Factors Behind Green Infrastructure

The efficacy of green infrastructure itself depends on various factors that exclude using their relationship with CO₂ as a gauge. Green infrastructure comes in many different forms and types, so there is a level of effectiveness that is dependent on the green infrastructure as a whole. There are a myriad of different factors that go into creating green infrastructure, ranging from the type of species use, the structure of the vegetation and the size, the proximity of the vegetation to sources of CO₂ emissions, as well as the management of the space itself by the city.⁹ Some forms of green infrastructure are simply more effective at dealing with CO₂ emissions than others. For example, while urban tree canopies and forests may do well with lowering temperatures and providing shade, they are not as effective with mitigating the emissions from vehicles on the roads. That is where vegetation closer to the ground, such as parks and roadside hedges and trees are more effective.^{9 3} Although it may seem obvious that different types of green infrastructure

have different applications when it comes to combating emissions, it is important to keep in mind to keep an accurate gauge of how effective a type of infrastructure will be.

Protection

The final measure of effectiveness when looking at green infrastructure is protection - measuring how well the vegetation protects city inhabitants from pollution and prevents contact as a whole. Typically, the goal of protection is achieved through the success in achieving the goals of reduction and extension, but it can be achieved on its own without the help of the others. However, it is worth noting that protection is the weakest form of combatance against emissions as it is far more effective to either reduce emissions or stop them at the source than let them out in the air and attempt to stop them down the line. Still, using green infrastructure as a form of protection does prove effective at offsetting CO₂ emissions by attempting to mitigate the effects, although this effectiveness is to a lesser extent than reduction and extension. One example of protection that can be used is the division of individuals from emission sources to prevent contact, seen through methods such as roadside hedges. These popular hedges are commonplace throughout large urban areas and serve as a natural barrier between pedestrians and vehicles on roads, forcing people to separate from the emissions sources, thus protecting them.³ The vegetation also has a further use of break down and absorbing emissions from cars, further reducing exposure of pedestrians to pollution while also accomplishing the previous two goals by reducing emissions by absorbing them and diluting them by breaking them apart. In attempting to protect individuals from emissions, green infrastructure is effectively able to offset the negative effects of CO₂ emissions by stopping the emissions from reaching city-dwellers and minimizing damage, while also accomplishing goals of further reducing and diluting emissions in the air.

Summary

When taking in this data and evidence, it is clear to see that green infrastructure is able to effectively offset the CO₂ emissions created in cities through its various methods. They are able to reduce emissions by employing large swaths of vegetation such as urban forests and rooftop gardens, absorbing CO₂ and directly reducing the emissions seen in cities. This decreases the amount of emissions in the air and leads to a healthier population and environment. Furthermore, the infrastructure is able to extend the distance that emissions travel through the use of natural barriers and infrastructure, such as roadside trees and urban forests, diluting CO₂ emissions in the air and making them less harmful while also making it easier for other forms of green infrastructure to further absorb the emissions. This improves the infrastructure's main purpose to reduce emissions, leading to a greater offset of emissions. Finally, the green infrastructure is able to accomplish its final goal of protection through the use of natural barriers such as hedges and bushes near population centers to block emissions from reaching inhabitants of cities outdoors, protecting them from the majority of the harmful effects that come with emissions. Additionally, their natural barriers serve as a further line of defense by also breaking apart emissions and

absorbing them, although to a lesser extent than their vegetation counterparts that are more adept to the task. When combined with the fact that green infrastructure is also able to effectively target the main sources of emissions as well as the effects of emissions - being heating, electricity, and transportation - it shows the effectiveness of green infrastructure when combatting emissions as they can offset the worst producers and effects of emissions and produce positive effects. While it is also true that effectiveness can be limited due to the differing factors and effects of the vast types of green infrastructure available, purposeful and clever management and implementation of green infrastructure would be able to bypass these limitations and allow the green infrastructure to operate at its maximum capacity and offset CO₂ emissions through varying and unique ways.

How would green infrastructure not offset CO₂ emissions?: A counter-argument and refuting evidence.

Although the benefits of green infrastructure are apparent in the data, there are those who argue that green infrastructure could pose a possible danger to cities and cause more harm than they do good through their methods. Green infrastructure can negatively impact the offsetting of CO₂ emissions due to the barriers surrounding them, such as the possible environmental obstacles of implementation, as well as the potential negative side-effects from implementing green infrastructure. Environmental struggles do exist when implementing green infrastructure, even though the main goal is to help the environment. Annually, global vegetation emits 0.5 to 1 billion tons of CO₂ into the atmosphere, leading to arguments against green infrastructure as vegetation in cities may put out more CO₂ than they take in, making the situation worse.⁹ Furthermore, the green infrastructure is unable to take in CO₂ at night as photosynthesis, the process used to pull CO₂ out of the air, is unable to activate without the sun. So, while the infrastructure sits idly, the emissions sources such as transportation and heating continue to run, not being offset by the vegetation. Furthermore, there are further concerns of damage to structures and urban infrastructure from green infrastructure. Tree and plant roots dig through concrete in search of nutrients in the soil, leading them to dig up sidewalks and roads, causing thousands of dollars in damages for the cities implementing them. This effect is exacerbated in green infrastructure such as green roofs and walls, where the roots and plants dig into the structure of the building, creating cracks and weakening the structural integrity.³ This can cause buildings to become dangerous as the added weight of the vegetation as well as the weakening structures can make buildings prone to collapse. This data leads individuals to argue against the implementation of green infrastructure, stating that the infrastructure poses a danger to the surrounding urban infrastructure and environment, doing more harm than good.

Although it is true that green infrastructure comes with certain negatives, it is important to see that the negatives are outweighed by the benefits of green infrastructure. While vegetation does produce CO₂, data has shown that the CO₂ produced by green infrastructure is less than the CO₂ it absorbs. A study conducted in Lugo, Spain showed that after green infrastructure was implemented in the city, there was a measured decrease in the CO₂ emissions in the city.⁵ Similar

patterns were observed in cities where green infrastructure was implemented, as CO₂ was absorbed and oxygen was released at a higher rate than the vegetation released CO₂. A notable example would be Chicago, with the Chicago Urban Forest Climate Project, a green infrastructure initiative that was able to remove 6,145 tons of air pollution over 3 years while isolating 155,000 tons of carbon annually.² This data contradicts the argument against green infrastructure as it shows that, contrary to the belief that green infrastructure could contribute to CO₂ emissions through the emitting of carbon, it is rather true that their absorption outweighs their output, making them effective at offsetting CO₂ emissions. Furthermore, on the point that green infrastructure poses a risk to buildings and urban infrastructure due to the growth of the vegetation and the spreading of roots, this claim is only as true as the city the green infrastructure is being built on makes it be. Regulations and careful planning from environmental engineers are put in place when constructing buildings with vegetation such as green roofs or other plants that could place a strain on a building. Buildings must account for that weight in their construction, otherwise the plan will not be approved by city officials. Additionally, incentives such as tax breaks or monetary incentives are given by cities to construction and architecture firms that incorporate green infrastructure into designs, ensuring that the benefits of green infrastructure are able to spread throughout the city with new developments. It has also been proven that the vegetation takes more and more CO₂ out of the air as it grows, showing that the growth of the plants and vegetation is not a downside, but rather a benefit that serves the city and its inhabitants.² This data and evidence demonstrates how green infrastructure provides more benefits to cities in offsetting CO₂ emissions rather than negatives, and that seeming negatives, such as the growth of roots in the ground, actually result in a net positive for cities and the combating of CO₂ emissions.

Conclusion

To summarize, it has been shown that there are a large variety of ways through which green infrastructure is effectively able to offset CO₂ emissions in cities. Analyzing the data, it is clear that there are multiple paths that green infrastructure takes in order to address emissions, often intertwining or overlapping in order to accomplish goals. Often, green infrastructure has been found to fall into one of three categories, reduction of emissions, extension of the distance emissions travel, and the protection of people or environments affected by emissions. In these categories, green infrastructure has proven to be very effective in accomplishing its purpose, being able to offset emissions by absorbing the emissions themselves through an abundance of vegetation, minimizing the damage caused by emissions by breaking them apart and diluting them with urban forests or rooftop gardens, or by simply separating and protecting those who could be affected with natural barriers like tall hedges or bushes.

Another route that green infrastructure can take is to target the main causes and effects of CO₂ emissions: transportation, electricity, and heating. By placing green infrastructure near sources, such as with trees near roads, or by directly combatting the effects of emissions, with urban canopies to provide natural shade and cooling, green infrastructure is once again able to

open a new avenue towards combatting and effectively offsetting emissions. Although limitations and arguments against the effectiveness of green infrastructure do exist, such as the potential damage of green infrastructure and the limited effectiveness of some types of vegetation, it has been proven that the benefits of green infrastructure far outweigh the potential negatives, and that cities are all the better for incorporating them into their design. Moving forward, it is important that cities become more conscious about the proven benefits of green infrastructure and take a larger initiative when it comes to incorporating it into designs. With cities being the largest producers of CO₂ emissions both today and in the projected future, it is necessary that those same cities take the first steps towards combating emissions and proactively incorporate green infrastructure into urban planning and design.

Acknowledgements

I would like to thank my mentor, Rachel Gehr of Perdue University for mentoring me through the writing process and providing support throughout the paper. I would also like to thank Lumiere Education for providing me with an opportunity to work with Rachel on my paper. Finally, I would like to thank my parents for supporting me in my research and learning endeavors during my writing process.

Works Cited

1. Mell, I.C. Green Infrastructure: concepts and planning. *FORUM Ejournal* [Online] **2008**, 8, 69-80. Google Scholar. <https://scholar.google.com/> (accessed Jan 18, 2024).
2. Molla, M.B. The Value of Urban Green Infrastructure and Its Environmental Response in Urban Ecosystem: A Literature Review. *CRDEEPJournals* [Online] **2015**, 4, 89-101. Google Scholar. <https://scholar.google.com/> (accessed Feb 23, 2024).
3. Hewitt N.C.; Ashworth K.; MacKenzie A.R. Using green infrastructure to improve urban air quality (GI4AQ). *SpringerLink* [Online] **2019**, 49, 62-73. Google Scholar. <https://scholar.google.com/> (accessed Jan 18, 2024).
4. Jayasooriya V.M.; Ng A.W.M.; Muthukumaran S.; Perera B.J.C. Green infrastructure practices for improvement of urban air quality. *Urban Forestry & Urban Greening* [Online] **2017**, 21, 34-47. ScienceDirect. <https://www.sciencedirect.com/> (accessed Feb 23, 2024).
5. De la Sota C.; Ruffato-Ferreira V.J.; Ruiz-García L.; Alvarez S. Urban green infrastructure as a strategy of climate change mitigation. A case study in northern Spain. *Urban Forestry & Urban Greening* [Online] **2019**, 40, 145-151. ScienceDirect. <https://www.sciencedirect.com/> (accessed Jan 18, 2024).
6. Glaesar E. L.; Kahn M.E. The greenness of cities: Carbon dioxide emissions and urban development. *Journal of Urban Economics* [Online] **2010**, 67, 404-418. ScienceDirect. <https://www.sciencedirect.com/> (accessed Feb 23, 2024).
7. Fragkakis M.; Lobo J.; Strumsky D.; Seto K.C. Does Size Matter? Scaling of CO₂ Emissions and U.S. Urban Areas. *Plos One* [Online] **2013**, e64727. Google Scholar. <https://scholar.google.com/> (accessed Feb 23, 2024).
8. Singh S.; Kennedy C. Environmental future energy use and CO₂ emissions of the world's cities. *Environmental Pollution* [Online] **2015**, 203, 271-278. ScienceDirect. <https://www.sciencedirect.com/> (accessed Feb 8, 2024).
9. Tallis M.J.; Amorim J.H.; Calfapietra C.; Freer-Smith P.; Kotthaus S.; Toscano P.; Lemes de Oliveira F.; Miranda A.I.; Grimmond S. The impacts of green infrastructure on air quality and temperature. *Social and Political Science 2015* [Online] **2015**, 30-49. Elgaronline. <https://www.elgaronline.com/> (accessed Feb 12, 2024).
10. Tiwari A.; Kumar P.; Kalaiarasan G.; Ottosen T. The impacts of existing and hypothetical green infrastructure scenarios on urban heat island formation. *Environmental Pollution* [Online] **2021**, 274, 115898. ScienceDirect. <https://www.sciencedirect.com/> (accessed Feb 23, 2024).
11. Kronenburg J.; Andersson E.; Barton D.N.; Borgström S.T.; Langemeyer J.; Björklund T.; Haase D.; Kennedy C.; Koprowska K.; Łaskiewicz E.; McPhearson T.; Stange E.E.; Wolff M. The thorny path towards greening: unintended consequence, trade-offs, and constraints in green and blue infrastructure planning, implementation, and management. *Ecology & Society* [Online] **2021**, 26, ES-12445-260236. Google Scholar. <https://scholar.google.com/> (accessed Feb 23, 2024).

12. Liu F.; Tait S.; Schellart A.; Mayfield M.; Boxall J. Reducing carbon emissions by integrating urban water systems and renewable energy sources at a community scale. *Renewable and Sustainable Energy Reviews* [Online] **2020**, *123*, 109767. ScienceDirect. <https://www.sciencedirect.com/> (accessed Mar 15, 2024).
13. Ai, K. Can Green Infrastructure Investment Reduce Urban Carbon Emissions: Empirical Evidence from China. *Land* [Online] **2024**, *13*, 226. MDPI. <https://www.mdpi.com/2073-445X/13/2/226> (accessed Feb 23, 2024).
14. Idiata, D. Understanding the Role of Green Infrastructure (GI) in Tackling Climate Change in Today's World. *International Journal of Environment and Sustainability* [Online] **2016**, *5*, 35-45. Science Target. <https://sciencetarget.com/1/> (accessed Feb 8, 2024).
15. Ramyar, R.; Zarghami, E. GREEN INFRASTRUCTURE CONTRIBUTION FOR CLIMATE CHANGE ADAPTATION IN URBAN LANDSCAPE CONTEXT. *Applied Ecology and Environmental Research* [Online] **2017**, *15*, 1193. Ebsco. <https://www.ebsco.com/> (accessed Jan 18, 2024).
16. Tóth, A.; Halajová, D.; Halaj P. GREEN INFRASTRUCTURE: A STRATEGIC TOOL FOR CLIMATE CHANGE MITIGATION IN URBAN ENVIRONMENTS. *Ecology & Safety* [Online] **2015**, *9*, 132-138. International Scientific Publications. <https://www.scientific-publications.net/en/> (accessed Feb 8, 2024).

To What Extent Does Public Opinion Impact Decisions Made in the Judicial System of The USA and the UK? By Shoni S.N. Doshi^{1*} Emmah K.S. Wabuke²

1. Ravenswood School For Girls, 10 Henry St, Gordon, Sydney, NSW, 2072
2. University of Cambridge, The Old Schools, Trinity Ln, Cambridge, UK, CB2 1TN

Abstract

The impact of public opinion has been prevalent in several current cases, including the Shamima Begum case and the Harvard affirmative action case. Previous research on these cases and the impacts of public opinion on the rulings has been extensive. However, my research paper will thoroughly analyze and conclude the extent of public opinion's effect on judicial decisions through a human rights lens that focuses on the qualitative aspects of the case's media influence and contextual influence, which other research papers may have yet to consider. First, I will set the context of the increasing significance of public opinion and then proceed with evaluating two cases in particular. This paper is going to make three arguments for each case. In Shamima Begum's case, I will first discuss the changes in her appearance. Secondly, I will highlight the influence of the terrorist attacks before this case. Lastly, I will examine the media's impact on public opinion. Regarding the Affirmative Action case, I will first analyze the social and racial dynamics that make affirmative action paradoxical.

Secondly, I will explore how societal dynamics have reigned over public opinion. Lastly, I will conclude how these aspects have subtly influenced the judicial discussions in both cases. To analyze and complete these arguments in my paper, I utilized multiple resources, including prior research papers and several articles from reliable news stations relating to the two cases I will discuss.

Keywords

Social Sciences; Public Opinion; Judicial System; USA; UK.

Introduction

The traditional interpretation of the law stems from a formal approach. This legal philosophy strongly emphasizes the strict and literal interpretation of legal rules and principles, often without considering the broader social, economic, and moral context in which these rules are applied.¹ While it is not the exclusive foundation of all legal proceedings, formalism significantly shapes the operations of legal systems in many jurisdictions.

Formalism holds significant importance in specific legal systems, particularly those prioritizing adherence to precedent and a strict interpretation of statutes and regulations. For instance, in common law systems like those in the United States and the United Kingdom, precedent (case law) and formal statutory interpretation are pivotal in legal decision-making. Formalism's influence is especially pronounced in areas of law where clarity, predictability, and consistency are highly valued, such as contract law or property law.

Nevertheless, the application of formalism can be complicated. Legal interpretation can vary significantly, as demonstrated by the cases discussed in this research paper. These cases illustrate how changes in public discourse can substantially impact the performance of legal principles and rules.

In recent years, public opinion has increasingly become a significant factor in legal proceedings,² shaping the outcomes of many high-profile cases. To what extent does public opinion impact the decisions made in the judicial systems of the USA and the UK? In this research paper, I will examine the impact of public opinion on judicial decision-making as a whole and then discuss its impact on two notable legal cases and how it has influenced the decisions made during and after the proceedings. These cases include the Shamima Begum case³ and the Harvard affirmative action case,⁴ of which have garnered significant media attention and sparked public debate.

Through an analysis of these cases, I will discuss how public opinion affects decision-making processes in the legal system and the extent to which judges and other key actors consider it. I will also explore the public battles surrounding each case, including the perspectives of those who championed different viewpoints and their impact on the legal outcomes. Overall, this research aims to shed light on the complex interplay between public opinion and the legal system, ultimately providing insights into how the law can better address the concerns and interests of the general public.

How Public Opinion Impacts Judicial Decision-making

In this paper, I define public opinion as “a means of informing decision-makers of the will of the people, especially towards public policy”.⁵

Public opinion can impact the process of judicial decision-making in several ways. Although bias should not be introduced into judicial decision-making, public opinion may still influence how legal actors, including judges and lawyers, interpret and apply the law to specific cases. Public opinion can also affect how judges and other legal decision-makers weigh the different factors involved in a case, such as the evidence presented and the interests of the parties involved.⁶ Furthermore, public opinion can shape the legal and political environment in which cases are decided. For example, public outrage or support for a particular issue or party can lead to changes in legislation or judicial appointments, which can, in turn, impact the outcome of future cases.

A great example of how public opinion has significantly impacted judicial decisions is the Kenyan LGBTQ+ NGO registration case.⁷ Public opinion plays a significant role in shaping Kenya’s government policies and actions. When it comes to LGBTQ+ issues, public opinion is often profoundly divided, with some segments of society holding conservative views that oppose LGBTQ+ rights. In contrast, others advocate for the rights of the LGBTQ+ community. In this case, The National Gay and Lesbian Human Rights Commission (NGLHRC) appealed to the Kenyan government for registration; however, their request was denied because their association with the LGBTQ+ community was “against the order of nature”.⁸

The government's controversial decision to deny the NGLHRC's application prompted a nationwide discussion. While some supported the decision, many expressed frustration by denouncing it as violating human rights.⁹ The supporters of the NGLHRC included LGBTQ+ activists and members of the LGBTQ+ community themselves. These supporters utilized social media to spread awareness about the issue leading to an uproar among the public.

Due to this social media uproar and the conflict within the country, international human rights organizations became aware of the public's outrage. They pressed the Kenyan government to reconsider the decision.¹⁰ In 2018, the High Court of Kenya found denying the NGLHRC's application unlawful; thus, the government was forced to register the organization.

Another fitting instance where the impacts of public opinion on the judicial decisions made in a case are clearly illustrated is the Johnny Depp and Amber Heard defamation case.¹¹ The court proceedings of this case occurred in Virginia, USA, and were televised and made public. Due to how high profile and debated the relationship between Amber Heard and Johnny Depp was before the case, the defamation trial became an increasingly popular topic of debate on social media and amongst other celebrities.

During the Amber Heard and Johnny Depp defamation trial, public opinion played a significant role in the judicial decisions. The intense media coverage surrounding the case may have affected the court's ability to remain impartial, especially since a jury was also involved. As the trial progressed, it became clear how divided the public had become. The videos of the trial were circulated through TikTok and Instagram, with captions from the posters, making their stance on the topic clear. Many TikTok users with large platforms even made skits of Amber Heard's testimonies to discredit her statements in court.¹²

Moreover, videos focused on the lawyer's mistakes on either side rather than the actual abuse and defamation discussed during the trial. In this case, it was clear how the public was divided towards Johnny Depp. This shift of public opinion to the side of a man in a case that revolves around domestic abuse has been seen by some as a public backlash of the #MeToo movement¹³ – a movement in which Amber Heard was a significant part. This overwhelming sympathy for Johnny Depp during the defamation trial is the public backlash due to the #MeToo movement, which shows how public opinion heavily influenced the outcome of the case.

Lastly, public opinion has recently been expressed and triggered by political statements that politicians themselves have made. The first politician that may come to mind in this scenario is Donald Trump – particularly the Muslim ban that he imposed in the US in 2017. Trump's tweets were often inflammatory and contributed to a highly polarized debate. During his campaign, Donald Trump openly discussed the possibility of banning Muslims from entering the United States. This rhetoric resonated with his supporters and catalyzed them to voice their support for such a ban online. Trump's followers saw his stance on immigration as an opportunity to advocate for stricter measures. As a result, they actively promoted the idea of a Muslim ban across various platforms.

With the overwhelming support of his voters and the influential position he held as the President of the United States, Trump used his platform to expand the ban beyond campaign

promises. This culminated in the implementation of immigration bans on citizens from 13 predominantly Muslim countries.

On the other hand, Trump's opponents saw the Muslim ban as discriminatory and unconstitutional. Trump's tweets, as well as general campaign promises, often emphasized discriminatory messages, with his critics accusing him of discriminating against a specific group of people based on their religion. Trump's opponents also argued that the ban was likely to do more harm than good, as it could aggravate tensions between the United States and Muslim-majority countries and potentially fuel anti-American sentiment. However, despite facing legal challenges from numerous immigrant rights organizations, the US Supreme Court ultimately allowed the ban to be implemented. Therefore, this suggests that the court's decision may have been influenced by the public opinion fostered by Trump and his ardent supporters, making it difficult for legal efforts to overturn the ban.

Overall, it is clear how public opinion can shape a case's political climate and thus influence its judicial decision-making. In this research paper, I will further this discussion and clarify to what extent public opinion influences judicial decision-making. The first case I will focus on is upholding the Home Secretary's decision to revoke Shamima Begum's citizenship, followed by the current case against Harvard by a group of Asian American students who claimed to have been denied admission from the university due to affirmative action.

How public opinion has impacted the outcome of the Shamima Begum case

In the following section of this paper, I will initially delve into the transformations in Shamima Begum's physical appearance. Next, I will emphasize the impact of the terrorist attacks that preceded this case. Finally, I will analyze the role of the media in shaping public perception, particularly in this case.

Shamima Begum was born and raised in the UK and was a British citizen at birth. She attended Bethnal Green Academy along with Amira Abase and Kadiza Sultan, two other girls who traveled with Shamima to Syria to join the Islamic State (IS). On February 17th, 2015, Shamima and her two friends traveled to Syria to join ISIS and become IS brides. Four years later, Shamima managed to escape the Islamic State successfully, but her two friends were killed in their attempt to escape. The Secretary of State halted Shamima's return to the UK for the Home Department, which deprived Shamima of her British citizenship because her return could pose a risk to national security. Since this decision was made, Shamima has remained in the Al-Roj camp run by the Syrian Democratic Forces. Moreover, Shamima's application for leave to enter on the basis that she wanted to appeal the decision of the deprivation of her citizenship was denied.

In this paper, I will be focusing on the Immigration Tribunal case that occurred in 2019, in which her application for leave to enter the UK was denied. Shamima Begum's case has been met with controversy and divided opinions. While some argue that she made a conscious decision to join ISIS and should not be allowed to return to the UK, others say that she was a

minor at the time and may have been coerced or groomed by ISIS recruiters. Furthermore, she has shown some remorse for her actions and desires to return to the UK to face justice. During her trial, Begum claimed that she was a victim and did not actively participate in any terrorist activities. However, some evidence suggests otherwise. While in Syria, she married an ISIS fighter and had three children, all of whom died due to illness while living in the war-torn region. The decision to strip Begum of her citizenship has also been criticized, with some arguing that it violates her human rights. Additionally, since she has been made stateless, the responsibility for her welfare and safety lies with Syria, a country amid a civil war.

This case is highly controversial in many aspects, and the reason for the heavy divide in public opinion between the general lay public and the public that is more knowledgeable about the law and legal practitioners themselves is why this case resulted in a considerably problematic decision. Although the lay public may think this case is the first of its kind, it is not. There have been several cases in the past where IS brides and even IS fighters were allowed to return to their country of citizenship to face prosecution, for instance, the case of Suhayra Aden,¹⁴ who was a Jihadi bride with a dual Australian and New Zealand citizenship. In this case, the Australian government revoked her citizenship so she would not be stateless. Instead, she is currently in a Turkish detention center awaiting prosecution.

Similarly, the United Kingdom's government is attempting to justify the revocation of Shamima Begum's citizenship because she also holds Bangladeshi citizenship. However, the point still stands that if she were to return to Bangladesh, she would ultimately face the death penalty. Nevertheless, several similar cases before the Shamima Begum case and the lack of legal and contextual knowledge amongst the lay public indicate the reason for the current uproar amongst the British people.

Additionally, there are a multitude of aspects, specifically regarding Shamima Begum herself, that have played a considerable role in shaping public bias. For example, how she is portrayed in the media, with alterations to her physical appearance and descriptions of her mental and behavioral characteristics, could have impacted the public's perception of her. Shamima Begum was commonly featured in Good Morning Britain wearing sports leggings, tank tops and snapbacks with her hair down. Shamima Begum first presented herself to the public during an interview with the Times Newspaper in 2019, where she was shown wearing an abaya and a hijab. This image of her being a Muslim woman who was not fazed after seeing decapitated heads in bins destroyed her portrayal in the public eye in the UK. However, during her interviews with Good Morning Britain, the image painted by her change in appearance was a simple teenage girl that the British public would consider normal to see on the street or the same clothes their daughters wear. This shift in appearance gave people the impression that Shamima Begum was trying to come across as innocent, childlike, and vulnerable and enhanced the argument that she was groomed as a child to join the Islamic State, a notion that Diane Abbott continues to share via her tweets and social media posts. However, the viewers of Good Morning Britain and the public think that the change in appearance was a futile attempt to portray herself

as normal and relatable. Begum claims that removing her hijab was for herself, but public perception refuses to believe her.

Moreover, in the Apple podcast “The Blame Game,” it is clear that the Muslim community – in which Shamima was involved before she joined the Islamic State – was extremely strict. Women, especially, were expected to wear hijabs in public and not listen to music, which they considered distractions. The fact that Shamima Begum is shown to be defying all of these rules on *Good Morning Britain* suggests that she has completely turned away from the radical ideology and mindset and has adopted a societally accepted attitude.

However, Begum’s interactions with the media and the public have had the opposite effect. In her interview with BBC News,¹⁵ even when talking about her two dead children, presented her as being inhumane and unrelatable. Viewers of *Good Morning Britain* claim that Shamima should not even be on camera that she has ultimately “disqualified herself from British life,” and that they should “leave her where she is and let her rot.”¹⁶ The anger of British viewers towards Begum’s interviews after her appearance had changed shows how the lay public is heavily influenced by her attitude towards the crimes committed by ISIS and towards the accusations claiming that she stitched people into suicide vests and took part in the Islamic State’s morality police.

Another primary reason why Begum’s perceived lack of remorse and attitude towards the accusations against her is due to the impact of the London bombings that occurred in 2005 and the Manchester bombing that occurred in 2017. The Manchester bombing was carried out by the Islamic State and resulted in 1017 casualties and 23 deaths in total. In her first interview with the *Times Newspaper*, she was asked about how she felt about the Manchester bombing and whether there was any reason for the attack. Although she claimed that innocent people were killed, she also claimed that there are “women and children in [ISIS territory] being killed right now unjustly with the bombings.” She added: “It’s a two-way thing really because women and children are being killed in the Islamic State right now, and it’s retaliation. Their justification was that it was retaliation, so I thought, OK, that is a fair justification.”¹⁷ These statements that insinuated that Shamima thought the Manchester bombing was justifiable left the public stunned and furious since the bombing only took place two years before Shamima Begum sought repatriation to Britain as well as the fact that the accusations towards her suggested that she could have been directly involved in the attack by stitching suicide bombers into their vests. The threat of ISIS after the Manchester bombing was felt amongst the British public, and Islamophobic attacks surged by 500%.¹⁸

The influence of these raucous attacks also continued to influence changes in terrorism laws in the UK. The main changes include the expansion of investigative powers, strengthening border security, and the Prevent strategy. The Investigatory Powers Act 2016, often called the “Snooper’s Charter,” broadened surveillance capabilities and granted authorities increased access to individuals’ communications and online activities in the interest of national security. The UK government also introduced more rigorous screening procedures at ports of entry, enhanced intelligence sharing with international partners, and implemented stricter visa regulations to

monitor and track individuals suspected of involvement in terrorism. However, the Prevent strategy was a highly problematic and debated strategy that significantly amplified anti-immigrant sentiment. The Prevent strategy aims to target areas seen as harbors for radicalization and terrorism. Mainly, Islamic cultural centers were isolated by Prevent due to the growing public sentiment against Muslim communities, especially after the London bombings and the Manchester Arena attack. Not only was the Prevent strategy influenced by public opinion, but it also managed to influence public opinion directly.

By targeting Islamic cultural centers, a direct target was put on any Muslim citizen of the UK. A fitting example of this is the Trojan Horse Affair. The Trojan Horse Affair began with a letter written to Birmingham's counsel, which named teachers and schools that supposedly radicalized students and went against British values. The actual effect of anti-immigrant sentiment that had stemmed from the previous terror attacks that Britain had faced was seen throughout the Trojan Horse Affair as the letter and its contents were investigated and debated in parliament even before the credibility of the letter was confirmed. It was found that the letter was written by a headmistress of a school who had lost a trial against the teachers mentioned in the letter and thus revealed that there was no truth to the claims made in the letter. The teachers involved in the Trojan Horse Affair faced significant career implications and difficulties securing future employment within the education sector due to the controversy and negative publicity surrounding their names. They faced public scrutiny and damage to their professional reputation as the media coverage of the affair brought their names into the public domain, often with negative connotations and assumptions that stimulated the anti-immigrant sentiment amongst the public.

Overall, it is clear that public opinion played a significant role in the Court of Appeals ruling upholding the Home Secretary's decision to revoke Shamima Begum's citizenship. The controversial nature of the case, coupled with the media portrayal of Begum as a victim or a terrorist, influenced public sentiment and created a divided public opinion on her return to the UK. The fear of terrorism and the impact of previous attacks further exacerbated anti-immigrant sentiment, which ultimately affected the immigration laws and policies in the UK. It is important to note that the lay public often overlooked the factual background and legal context of the case and that the decision to revoke Begum's citizenship was primarily based on the risk to national security. Therefore, it is clear how public opinion painted an unfavorable image of Shamima Begum on her behalf, which many see impacted the decision to uphold the revocation of her citizenship.

How public opinion has impacted the judicial discussions in the US Supreme Court regarding affirmative action

In this section of the paper, I will initially examine the paradoxical nature of affirmative action and discuss the prior cases that revolve around quotas surrounding diversity, focusing on the intertwined social and racial dynamics. I will then investigate how prevailing societal

dynamics have shaped public opinion. In conclusion, I will assess how all these elements have subtly influenced the judicial deliberations in both cases.

Affirmative action is an excellent example of how public opinion has impacted judicial attitudes in the United States. It is a policy implemented to redress historical and ongoing discrimination and has long been a subject of intense debate and legal scrutiny in the United States. In this research paper, I will define affirmative action as a set of practices undertaken by employers, university admissions offices, and government agencies to go beyond non-discrimination to actively improve the economic status of minorities regarding employment, education, and business ownership and growth.¹⁹ Over the years, numerous legal challenges have emerged, challenging the constitutionality and fairness of race-conscious admission policies. One prominent case I will focus on in this research paper, *Students for Fair Admissions v. Harvard*,²⁰ has garnered significant attention and become a focal point in affirmative action discourse. This case, filed in 2014 by the Students for Fair Admissions organization, alleges that Harvard University's admissions process unfairly disadvantages Asian American applicants through racial preferences.

The main aspects of the case that I will be focusing on in this research paper are, firstly, the discrimination of certain ethnic groups to benefit other ethnic groups; secondly, the sudden poise of the US Supreme Court towards the overturning of affirmative action; and lastly, the prevalent hierarchy of race and how public opinion has impacted all of these aspects. The reason why this case is highly controversial is because it highlights to the public the actual ramifications of affirmative action on other ethnic minorities in the US. Moreover, this case has become the forefront of debates regarding college admission procedures. The main reason for introducing affirmative action was to address historical inequalities and promote diversity in educational institutions.

Throughout the 1970s, the U.S. Supreme Court was at the forefront of reflecting and shaping public opinion on affirmative action in university admissions. *Regents of the University of California v. Bakke (1978)*²¹ was pivotal. In this case, the Court ruled that while racial quotas were unconstitutional, considering race as one factor among many in university admissions to promote diversity was constitutionally permissible. This decision underscored the evolving and contentious nature of affirmative action policy, as it sought to strike a balance between addressing historical racial inequalities while simultaneously avoiding reverse discrimination. During this period, the substantial support for these policies was driven by the civil rights movement and a recognition of the need to redress historical discrimination. However, as the 1980s progressed, opposition to affirmative action grew, fueled by concerns of reverse discrimination, particularly amongst Caucasian Americans.

The demographic changes in the United States played a substantial role in shaping public opinion. The country became increasingly diverse due to immigration from Asia, Latin America, and other regions, in addition to the growing presence of African Americans. The influx of new ethnic and racial minority groups introduced new dimensions to the affirmative action debate.

Critics argued that affirmative action should consider broader factors, including socioeconomic status, to address inequalities more effectively.

One case that reflected these evolving demographics was the *City of Richmond v. J.A. Croson Co. (1989)*²². In this case, the Court struck down a municipal affirmative action plan that set aside a percentage of contracts for minority-owned businesses. The decision highlighted the need for a compelling state interest and narrowly tailored measures when implementing affirmative action. It also demonstrated the Court's efforts to balance the interests of a diverse society.

In subsequent decades, the United States continued to experience demographic changes with the growth of Hispanic, Asian, and multiracial populations. Public opinion on affirmative action remained dynamic and complex, influenced by race, political affiliation, and personal experiences. This ongoing evolution led to further Supreme Court cases such as *Grutter v. Bollinger (2003)*²³ and *Fisher v. University of Texas at Austin (2013)*,²⁴ which upheld the idea of diversity as a compelling interest in higher education but required institutions to use race-conscious policies with strict scrutiny and limited scope.

In the seventies and eighties, the US Supreme Court was liberal-leaning and mainly favored affirmative action in college admissions. The court opposed public opinion within the US at the time, although it was the height of the civil rights movement, the White elitist community being the majority. However, the Supreme Court's ideological composition shift has been distinct after Trump's presidency.

From a liberal-leaning Court in the seventies and eighties to a more conservative Court due to appointments by Donald Trump, many conservative Justices have expressed their opposition to affirmative action, which has influenced the Court's stance on the matter and allowed other marginalized groups to be heard such as the Asian American community who have been leading this case.

Additionally, the changing landscape of political correctness and the rise of racial justice movements, such as the Black Lives Matter (BLM) movement and the election of Barack Obama, has brought racial inequalities to the foreground of public discourse. However, these developments have often centered on the experiences of African Americans, potentially leaving the concerns of other marginalized groups unaddressed.

BLM raised awareness of racial discrimination, police brutality, and systemic racism, but its primary focus was on issues affecting the Black community. Similarly, Barack Obama's historic presidency symbolized progress, but it was primarily viewed as a milestone for African Americans. While his administration addressed racial justice, it often emphasized the Black experience over other marginalized groups, sparking debates about the breadth of progress regarding diversity across the United States of America.

Public opinion shifted as more individuals recognized racial disparities. Still, this shift often revolved around challenges facing African Americans, limiting the space for the concerns of other ethnic minority groups to be acknowledged. This dynamic underscores the need for a

more inclusive approach to addressing racial inequalities, considering the experiences of all marginalized communities.

Affirmative action supporters believe that considering factors such as race or ethnicity in admissions can help level the playing field and provide opportunities for underrepresented groups. However, the counterargument that is becoming increasingly prevalent is how only specific ethnic minorities and underrepresented groups are being aided by affirmative action and that affirmative action is leaving ethnic minorities such as Asian Americans out of the equation. Moreover, critics of affirmative action argue that it can lead to reverse discrimination or the unfair exclusion of individuals based on their race or ethnicity. They contend that admissions decisions should be strictly merit-based and that considering race as a factor can violate the principle of equal treatment under the law. A broader social, political, and ideological perspective amongst the public has explicitly influenced this case.

The PEW research center's public polls²⁵ show how approximately 69% of those surveyed believe that race should not be considered a factor in college admissions, and only an average of 10% thought that race should be a significant factor. This indicates that most Americans believe affirmative action should not affect college admissions. This rift in the public is also reflected in the US Supreme Court. This division in public opinion is reflected in the US Supreme Court, where six Justices are conservatives, and three are democrats, implying a majority rule against affirmative action. Furthermore, the Supreme Court's Justices have reflected the public majority throughout the case by suggesting race-neutral alternatives in college admissions. However, the University of California and the University of Michigan both claimed that race-neutral options cost millions of dollars and are significantly less efficient and effective and instead result in fewer minority students.²⁶ These claims led to a public backlash as affirmative action opponents claimed that the alternatives were necessary and that colleges should consider the cultural experiences of individuals rather than just their race to ensure diversity.

In contrast, many members of the American public needed help to identify a clear distinction between culture and race. This debate was mirrored in the Supreme Court as Patrick Strawbridge, a Student for Fair Admissions representative, suggested how cultural experiences instead of race can be utilized in college admissions. Justice Kagan rebutted this notion, claiming that "race is a part of the culture and culture is part of the race".²⁷ Justice Kagan's statement reflects the vast divide in public opinion, especially within the Asian American community. Certain cases have been alleged to have discriminated against Asian American applicants by affirmative action policies in college admissions. One such case is the Harvard affirmative action case, which I am currently discussing. This lawsuit claimed that Harvard discriminates against Asian American by giving them lower scores on the personal rating,²⁸ which measures subjective factors such as unique qualities and character. The lawsuit alleged that this practice resulted in a lower acceptance rate for Asian American applicants compared to other racial groups.

The implementation of affirmative action has created unintended consequences and resulted in bias towards certain groups, including Asian Americans, who are often perceived as overrepresented in higher education institutions. Furthermore, some studies have shown that Asian Americans need higher test scores and academic achievements to gain admission to selective universities compared to other groups. This suggests that affirmative action policy may limit opportunities for high-achieving Asian American students, particularly in competitive institutions.

The debate surrounding Justice Kagan's statement is complex and ingrained in all minority communities. Yet, through affirmative action, only selected minorities are seen as having cultural backgrounds that could hinder them, while other minorities are caught in an opposite light. A clear example of the overlap of culture and race is the model minority myth that surrounds the Asian community and culture in general. The model minority myth perpetuates the stereotypical narrative that Asian Americans are academically and musically inclined and, therefore, are a model minority. The main disadvantage of the model minority myth is that it masks the reality of Asian Americans who face economic and social obstacles that prevent them from accessing the resources and opportunities they need – like the Hispanic and African American communities. The public opinion of, specifically, the Asian American community, according to a study conducted by PEW Research Centre, is that Asian Americans have the highest poverty rate among any racial or ethnic group in New York City despite being viewed as high achievers.²⁹ This model minority myth, paired with the blatant discrimination of affirmative action, has worked against the Asian American community and has disadvantaged them even in an area they supposedly excel in. The solidarity between the Asian American community and the African-American community has now become notably strained due to affirmative action. The history of solidarity between both minority communities has spanned since the Civil Rights Movement, which is evident since the term "Asian American" was invented at the height of the movement. It has been shown that the Asian American community is victim to the hierarchy of race in which African Americans gain the most advantage amongst other minorities in the USA. This is mainly due to the systemic racism and past sufferings of the African-American community, and affirmative action was seen as a step taken to right these past wrongs that African Americans had faced. This links with how Peter Boghossian mentioned in an interview with Fox News that colleges were "looking for quick answers to difficult problems, and that is an injustice to everybody"³⁰ regarding implementing affirmative action. However, the main argument the Asian American community is expressing is about why affirmative action is seen as a way of righting the wrongs of minority groups when not all minority groups benefit from affirmative action.

In the wake of the Supreme Court's decision to ban affirmative action, the justices' opinions have shed light on the complexities and implications of this pivotal ruling. Chief Justice John Roberts, writing for the majority, emphasized the importance of treating applicants as individuals, emphasizing that they can still share their racial experiences in college applications.

However, the crux of the majority opinion lies in the idea that race should not be the decisive factor in the admissions process, prompting universities to consider applicants holistically. Conversely, the dissenting Democratic-appointed justices strenuously objected to this decision, arguing that it would virtually preclude tertiary institutions from factoring race into their admissions processes. Justice Sonia Sotomayor underscored the inescapable reality that race remains a significant factor in students' lives. She criticized the majority's stance as a "false promise," asserting that the decision attempts to save face and appears detached from the prevailing realities of racial disparities.

When viewed in its entirety, the Supreme Court's ruling represents a substantial change in university admissions. It prohibits universities from using race as a decisive factor. Still, it allows consideration of other elements, such as cultural experiences and the challenges associated with an applicant's racial background, as evidenced in college essays. These alternative avenues enable universities to continue pursuing diversity on their campuses, which was the fundamental aim of affirmative action.

However, the Supreme Court's decision to ban affirmative action exempts the US military service academies. This exemption is rooted in recognizing "potentially distinct interests" between the two types of institutions articulated by the majority opinion. One core differentiating factor lies in the necessity for representation within military leadership, an imperative inherent to service academies' unique mission and function. Unlike general universities, which draw students from around the world regardless of citizenship, the military academies exclusively enroll US citizens. Consequently, they are mandated to be representative of the broader US demographic, reflecting the nation's diversity.

This contrast is key to understanding the Court's decision. The ruling acknowledges that universities' primary objective is not to mirror the US demographic, as it is with military academies, but rather to create diverse, inclusive, and global learning environments. As such, the ban on affirmative action in university admissions aims to ensure that race is not used as a basis for discrimination that exclusively benefits one particular racial group, such as the African-American population within the United States.

Moreover, it is important to note that service academy graduates constitute a minority, making up fewer than 20 percent of all military officers. This further underscores the distinctive nature of the military academies and their unique responsibility to represent the broader US population in their officer corps.

In essence, the Supreme Court's decision brings into sharp relief the delicate balance between acknowledging the role of race in individual experiences and striving for fairness and equality in university admissions while also recognizing the distinctions between the US military academies and regular tertiary institutions. While it marks a significant departure from the affirmative action policies of the past, it opens the door to new means of fostering diversity in higher education, demonstrating the ongoing complexity of this critical issue. The ruling prompts us to confront the profound challenges of promoting inclusivity while respecting the principle of treating individuals as unique, multifaceted beings.

To surmise, the Harvard affirmative action case has highlighted the contentious issues surrounding the role of affirmative action in college admissions and its impact on various ethnic groups. The case has revealed how affirmative action policies can be discriminatory towards specific groups, such as Asian Americans, and how it may limit opportunities for high-achieving students. Public opinion has played a significant role in the debate surrounding affirmative action, with a clear divide between those who support race-conscious admissions and those who oppose it. The US Supreme Court has reflected this divide regarding this particular case, with a conservative majority ruling against affirmative action, although having opposed public opinion in prior cases in the late seventies and early eighties. The case has also highlighted the hierarchy of race in the US, with African Americans benefiting the most from affirmative action and other minority groups, such as Asian Americans, being left out. This has led to a strained relationship between minority communities and has challenged the notion of solidarity between them. In conclusion, the Harvard affirmative action case has demonstrated how public opinion can impact federal rulings and how affirmative action remains a highly contentious issue in the US today.

Conclusion

To conclude the entire paper, both the Shamima Begum case and the Harvard affirmative action case have shown the significant impact of public opinion on legal decisions. In the Begum case, public sentiment played a role in her citizenship being revoked, despite the legal context and risk to national security being the primary factors. In the Harvard case, public opinion has influenced the debate surrounding affirmative action, highlighting the divide between those who support race-conscious admissions and those who oppose it. These cases reveal the importance of public education and understanding legal issues beyond media portrayals and political rhetoric.

Acknowledgments

Thank you for the guidance of my mentor, Emmah Khisa Senge Wabuke, from the University of Cambridge, in developing this research paper.

Works Cited

- Nishigai, K. (2021) Two Types of Formalism of the Rule of Law, Academic.oup.com.
<https://academic.oup.com/ojls/article/42/2/495/6420390> 02 October 2023.
- Voeten, Erik. "Public Opinion and the Legitimacy of International Courts." *Theoretical Inquiries in Law*, vol. 14, no. 2, Jan. 2013, <https://doi.org/10.1515/til-2013-021>.
- R (on the application of Begum) v Special Immigration Appeals Commission ([2021] UKSC 7)
- Students for Fair Admissions, Inc. v. President and Fellows of Harvard College ([2022] 20-1199)
- Sean D. Carey *The Concise Oxford Dictionary of Politics*
- Mason, Anthony. *The Courts and Public Opinion*. National Institute of Government and Law's inaugural public lecture on "The Courts and Public Opinion"
- NGOs Co-ordination Board v EG & 4 others [2023] KESC 17 (KLR)
- Gathara, Patrick. "How an LGBTQ Court Ruling Sent Kenya into a Moral Panic." *Www.aljazeera.com*,
www.aljazeera.com/opinions/2023/3/15/how-an-lgbtq-court-ruling-sent-kenya-into-a-moral-panic.
- Mersie, Ayenat. "For LGBTQ Kenyans, Court Win Prompts Backlash as Threats Escalate." *Reuters*, 20 Apr. 2023,
www.reuters.com/world/africa/lgbtq-kenyans-court-win-prompts-backlash-threats-escalate-2023-04-18/.
- "Kenya Wrong to Ban LGBT Rights Groups from Registering - Supreme Court." *BBC News*, 24 Feb. 2023, www.bbc.com/news/world-africa-64491276.
- John C. Depp, II v. Amber Laura Heard
- Goldberg, Michelle. "Opinion | Amber Heard and the Death of #MeToo." *The New York Times*, 18 May 2022, www.nytimes.com/2022/05/18/opinion/amber-heard-metoo.html.
- Grady, Constance. "Johnny Depp's Legal Victory Makes It Clear the Me Too Backlash Has Arrived." *Vox*, 2 June 2022,
www.vox.com/culture/23150632/johnny-depp-amber-heard-trial-verdict-me-too-backlash
- Aden v Minister of Internal Affairs [2021] NZHC 446
- "“I Didn't Want to Be IS Poster Girl.”" *BBC News*, 18 Feb. 2019,
www.bbc.com/news/uk-47276572.
- Good Morning Britain. "Shamima Begum Offers to Help Boris Johnson in Fight against Terrorism If She Can Return to UK | GMB." *YouTube*, 15 Sept. 2021,
www.youtube.com/watch?v=WtI_CKDvn9Y. Accessed 12 March. 2023.
- Dearden, Lizzie. "Shamima Begum: Manchester Arena Bombing "Justified" because of Syria Airstrikes, Isis Teenager Says." *The Independent*, 18 Feb. 2019,
www.independent.co.uk/news/world/middle-east/shamima-begum-isis-interview-manchester-bombing-terror-attack-syria-airstrikes-a8784741.html.
- Halliday, Josh. "Islamophobic Attacks in Manchester Surge by 500% after Arena Attack." *The Guardian*, *The Guardian*, 22 June 2017,

www.theguardian.com/uk-news/2017/jun/22/islamophobic-attacks-manchester-increase-arena-attack.

Harry J. Holzer, *Affirmative Action: What Do We Know?* 2005.

Students for Fair Admissions, Inc. v. President and Fellows of Harvard College ([2022] 20-1199)

Regents of Univ. of California v. Bakke, 438 U.S. 265 (1978)

City of Richmond v. J. A. Croson Co., 488 U.S. 469 (1989)

Grutter v. Bollinger, 539 U.S. 306 (2003)

U.S. Reports: *Fisher v. University of Tex. at Austin*, 570 U.S. 297 (2013)

Graf, Nikki. “Most Americans Say Colleges Should Not Consider Race or Ethnicity in Admissions.” Pew Research Center, 25 Feb. 2019, www.pewresearch.org/short-reads/2019/02/25/most-americans-say-colleges-should-not-consider-race-or-ethnicity-in-admissions/.

Ax, Joseph. “Explainer: What Happens If the US Supreme Court Bans Affirmative Action?” Reuters, 27 Mar. 2023, www.reuters.com/legal/what-happens-if-us-supreme-court-bans-affirmative-action-2023-03-27/.

Sneed, Ariane de Vogue, Tierney. “Takeaways from SCOTUS Affirmative Action Cases: Conservatives May Overturn Precedent Allowing Race as a Factor in Admissions | CNN Politics.” CNN, 31 Oct. 2022, edition.cnn.com/2022/10/31/politics/takeaways-supreme-court-harvard-north-carolina-affirmative-action/index.html. Accessed 3 Sept. 2023.

Kaur, Harmeet. “How Asian Americans Fit into the Affirmative Action Debate.” CNN, 3 Nov. 2022, edition.cnn.com/2022/11/03/us/affirmative-action-asian-americans-qa-cec/index.html.

Pew Research Center. “The Rise of Asian Americans.” Pew Research Center’s Social & Demographic Trends Project, 19 June 2012, www.pewresearch.org/social-trends/2012/06/19/the-rise-of-asian-americans/.

Reyes, Gabrielle. “Professor Says This University Practice Is ‘Profoundly Un-American.’” Fox News, 13 Jan. 2023, www.foxnews.com/politics/peter-boghossian-affirmative-action-college-admissions-profoundly-un-american.

Navigating the Digital Shift: The Impact of E-Commerce on Brick-and-Mortar Retail Strategies and Consumer Behavior By Rebecca Blumberg

Introduction

E-commerce, defined as the buying and selling of goods and services over the internet, has undergone tremendous evolution since its inception. The first recorded online transaction occurred in 1994 when a Sting album was sold through a website called NetMarket. This marked the beginning of a new era in retailing, leading to the development of sophisticated online marketplaces. The late 1990s saw the emergence of major e-commerce players like Amazon (founded in 1994) and eBay (founded in 1995), which revolutionized the way consumers shopped by offering a vast selection of products and competitive pricing. According to a report by the National Bureau of Economic Research (NBER), e-commerce sales in the United States grew from virtually zero in the mid-1990s to over \$100 billion by the early 2000s. The growth trajectory of e-commerce continued to accelerate with advancements in technology. The widespread adoption of broadband internet in the early 2000s made online shopping more accessible.

The introduction of smartphones and mobile internet in the late 2000s and early 2010s further propelled e-commerce growth, enabling consumers to shop anytime and anywhere. A report by McKinsey & Company highlights that mobile commerce (m-commerce) has become a significant segment of e-commerce, accounting for nearly 45% of total e-commerce sales by 2020. The COVID-19 pandemic in 2020 was a pivotal moment for e-commerce. With lockdowns and social distancing measures in place, consumers turned to online shopping in unprecedented numbers. According to Business Insider, global e-commerce sales surged by 27.6% in 2020, reaching \$4.28 trillion. As shown in Figure 1, this rapid growth highlighted the critical role of e-commerce in the modern retail landscape and underscored the need for traditional retailers to adapt to the changing consumer behaviors. I want to write this paper because the rapid growth of e-commerce fascinates me, particularly how it has transformed traditional retail operations and consumer behavior. I'm intrigued by the innovative strategies brick-and-mortar retailers are adopting to stay competitive in a digital-first world, and I want to explore the broader implications of these changes on industries and local economies.

Understanding the dynamic interplay between online and offline shopping experiences is crucial in today's economy, and I'm eager to delve into how these shifts are shaping the future of retail. The rapid growth of e-commerce has compelled brick-and-mortar retail companies to adopt new operational strategies, including the integration of omnichannel approaches and technological adaptations. These changes have significantly impacted consumer shopping behavior, reshaped industries, and altered consumer preferences. This paper will explore the growth and trends of e-commerce, the operational strategies brick-and-mortar retailers have implemented in response, and the subsequent effects on consumer behavior and industry dynamics.

CONSUMPTION SPREADS FASTER TODAY

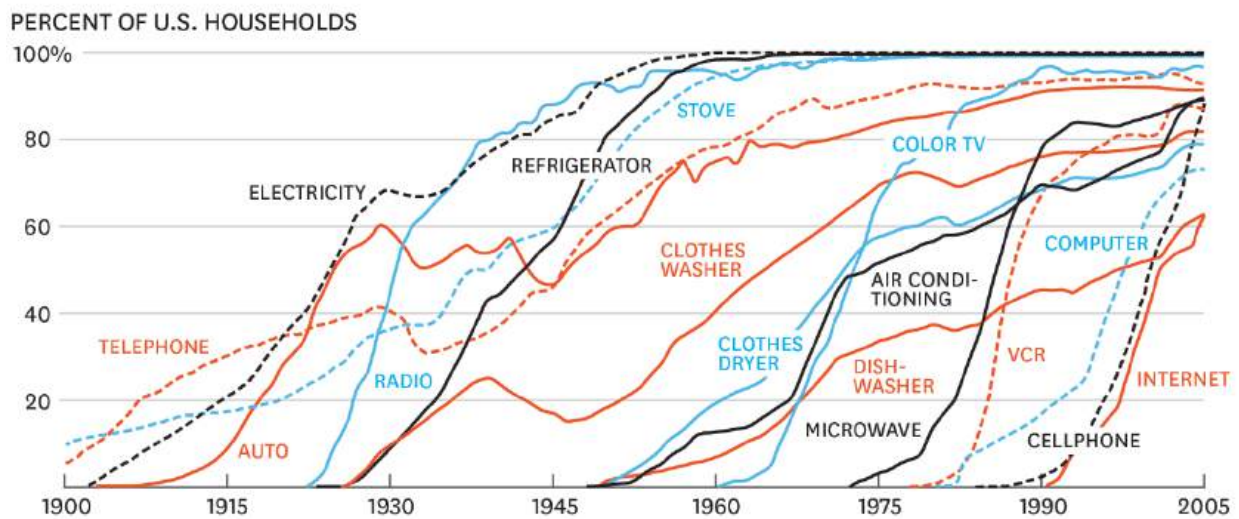


Figure 1

The History of E-Commerce and its Changing Landscape

Before the rise of e-commerce, brick-and-mortar retail was the predominant model for consumer shopping. Traditional retail stores have a long history, with department stores like Macy's and Sears becoming household names in the early 20th century. These stores offered a wide range of products under one roof, providing consumers with convenience and a variety of choices. In the post-World War II era, the suburbanization of America led to the rise of shopping malls, which became social and commercial hubs. By the 1980s, malls and big-box retailers like Walmart and Target dominated the retail landscape. These stores focused on offering a vast selection of products at competitive prices, often leveraging economies of scale to attract price-sensitive consumers. Despite the success of these retail formats, they faced several challenges. High operating costs, including rent, utilities, and labor, made profitability a constant concern. Additionally, traditional retailers had to manage complex supply chains and inventory systems to ensure that products were available to meet consumer demand. According to a study by Cushman & Wakefield, brick-and-mortar retailers also faced significant competition from each other, leading to price wars and thin profit margins. The arrival of e-commerce introduced a new set of challenges for brick-and-mortar retailers. The convenience of online shopping, coupled with the ability to compare prices and read reviews, appealed to consumers. As e-commerce platforms like Amazon offered fast shipping and easy returns, traditional retailers struggled to keep pace. The rise of digital technology and shifting consumer preferences forced brick-and-mortar retailers to reevaluate their operational strategies and explore new ways to stay relevant in an increasingly digital world.

E-commerce has experienced exponential growth over the past few decades, with sales increasing from nearly zero in the mid-1990s to over \$4 trillion globally by 2020. Key

milestones in this growth include the establishment of major platforms like Amazon and eBay in the mid-1990s, the widespread adoption of broadband internet in the early 2000s, and the advent of smartphones and mobile internet in the late 2000s and early 2010s. According to Business Insider, global e-commerce sales surged by 27.6% in 2020 alone, driven by the COVID-19 pandemic and the subsequent shift in consumer behavior towards online shopping. Technological advancements such as secure payment systems, user-friendly interfaces, and sophisticated logistics networks have further fueled this growth, making online shopping more accessible and convenient for consumers worldwide. Several factors have driven the consumer shift to online shopping. Increased internet penetration and the proliferation of mobile technology have made it easier for consumers to access online stores from anywhere at any time. A report by Cushman & Wakefield indicates that mobile commerce (m-commerce) has become a significant segment of e-commerce, accounting for nearly half of all e-commerce sales by 2020. Additionally, the convenience of home delivery, the ability to compare prices easily, and the availability of a wider range of products online have made e-commerce an attractive option for many consumers. The COVID-19 pandemic further accelerated this shift, as lockdowns and social distancing measures forced consumers to turn to online shopping for their needs. According to ScienceDirect, the impact of these factors has been profound, leading to a permanent change in consumer shopping behavior and expectations.

Changing Strategies & New Technologies

To compete with the convenience and reach of e-commerce, many brick-and-mortar retailers have adopted omnichannel retailing strategies. Omnichannel retailing involves integrating various shopping channels, including physical stores, online platforms, and mobile apps, to provide a seamless and cohesive shopping experience for consumers. Successful case studies, such as those of Walmart and Target, demonstrate how these retailers have leveraged their physical stores as fulfillment centers for online orders, offering services like buy online, pick up in-store (BOPIS), and curbside pickup. ResearchGate highlights that these strategies not only enhance customer convenience but also increase store traffic and sales. Brick-and-mortar retailers have also embraced digital tools and platforms to enhance their operational efficiency and customer experience. This includes implementing advanced inventory management systems, data analytics, and customer relationship management (CRM) software. According to Springer, these technologies enable retailers to optimize their supply chains, reduce costs, and better understand consumer preferences and behaviors. Additionally, retailers have introduced in-store technologies such as augmented reality (AR) for virtual try-ons, mobile checkout systems, and personalized recommendations to attract and retain customers. To compete with the convenience of online shopping, brick-and-mortar retailers have focused on enhancing the in-store experience. This includes redesigning store layouts to make shopping more enjoyable and interactive, incorporating technology such as AR and mobile checkout to streamline the shopping process, and offering personalized services and experiences that cannot be replicated online. Business Insider reports that retailers like Sephora and Apple have successfully used

technology to create unique in-store experiences, attracting customers and building brand loyalty. (Sources: JTAER, ResearchGate)

Brick-and-mortar retailers have increasingly adopted digital tools and platforms to streamline operations and enhance customer experiences. One significant adaptation is the implementation of advanced inventory management systems, which use data analytics and machine learning to optimize stock levels, predict demand, and reduce waste. According to Springer, these systems enable retailers to maintain the right balance of inventory, ensuring that popular items are always in stock while minimizing excess inventory costs. Additionally, retailers have integrated customer relationship management (CRM) software to better understand consumer behaviors and preferences, allowing for personalized marketing and improved customer service. McKinsey & Company highlights that the adoption of digital platforms, such as e-commerce websites and mobile apps, has allowed traditional retailers to expand their reach, offering customers the convenience of online shopping while leveraging their physical stores for services like same-day delivery and in-store pickups. Enhancements in supply chain and inventory management have been crucial for brick-and-mortar retailers adapting to the rise of e-commerce. The implementation of automated systems and robotics in warehouses has significantly improved efficiency and accuracy in order fulfillment. According to McKinsey & Company, these technologies enable faster processing of online orders and reduce the likelihood of errors, leading to higher customer satisfaction. Additionally, real-time tracking and data analytics have enhanced supply chain visibility, allowing retailers to respond quickly to changes in demand and potential disruptions. This agility is particularly important in maintaining a competitive edge in the fast-paced retail environment. (Sources: Springer, McKinsey & Company)

To compete with the convenience of online shopping, brick-and-mortar retailers have focused on enhancing the in-store experience through the use of technology. Augmented reality (AR) and virtual reality (VR) are being used to create interactive and immersive shopping experiences. For example, retailers like Sephora use AR technology to allow customers to try on makeup virtually, enhancing the shopping experience without the need for physical samples. Mobile checkout systems have also been implemented to reduce wait times and improve customer convenience. According to ScienceDirect, these technologies not only enhance the customer experience but also provide valuable data that retailers can use to further personalize services and offerings. Changes in store layout and services have also been instrumental in attracting customers. Retailers are redesigning stores to create more engaging and interactive environments. For instance, Apple stores are known for their open, minimalist layouts that encourage customers to explore and interact with products. Business Insider reports that retailers are incorporating experiential elements such as product demonstrations, workshops, and events to draw customers into physical stores. These enhancements aim to provide unique experiences that cannot be replicated online, fostering customer loyalty and driving foot traffic. (Sources: ScienceDirect, Business Insider) The rise of e-commerce has led to a significant shift in consumer purchasing patterns, with more people opting to shop online rather than visit

physical stores. This shift has resulted in decreased foot traffic for brick-and-mortar retailers, as consumers prioritize the convenience and variety offered by online platforms. Studies from sources like Finatic and CNU highlight that consumers increasingly value the ability to shop at any time, access a broader range of products, and easily compare prices across different retailers. This change in behavior has forced traditional retailers to rethink their strategies, often leading to the adoption of omnichannel approaches that blend the benefits of online and in-store shopping. In response to the growing popularity of e-commerce, consumer expectations and demands have also evolved. There is now an increased demand for faster delivery times, personalized shopping experiences, and seamless transactions across multiple channels. Social media and online reviews play a crucial role in influencing consumer decision-making, as shoppers rely on these platforms to gather information and recommendations before making purchases. Cushman & Wakefield and Finatic report that retailers must adapt to these new demands by enhancing their logistics capabilities, offering tailored marketing efforts, and maintaining a strong online presence. The pressure to meet these heightened expectations has led to significant changes in how retailers approach customer service and product offerings.

Increased Competition & Evolving Consumer Preferences

The impact of e-commerce on the competitive landscape of the retail market is profound. E-commerce giants like Amazon and Alibaba have reshaped the industry by leveraging their extensive logistics networks, vast product selections, and sophisticated data analytics to dominate the market. These companies have set new standards for speed, convenience, and personalization, forcing traditional retailers to innovate and adapt. In an effort to compete, many brick-and-mortar retailers have adopted strategies such as expanding their online presence, enhancing their delivery options, and offering exclusive in-store experiences. ResearchGate and HilariPublisher provide examples of how traditional retailers are utilizing their physical stores as fulfillment centers for online orders and integrating digital tools to streamline operations. The economic impacts of e-commerce on the retail sector are wide-ranging. On one hand, the shift towards online shopping has led to job losses in traditional retail roles, particularly in physical stores that have seen declining sales. On the other hand, e-commerce has created new opportunities in areas such as logistics, warehousing, and digital marketing. The National Bureau of Economic Research (NBER) and McKinsey & Company highlight that investment and capital allocation within the retail sector have shifted towards technology and digital infrastructure, as companies prioritize the development of robust e-commerce platforms and the integration of advanced technologies. These changes have had significant effects on local economies, as the decline of brick-and-mortar stores has impacted commercial real estate markets and reduced tax revenues for municipalities dependent on retail sales.

Consumer preferences in the retail sector have evolved significantly with the rise of e-commerce. Today, consumers prioritize convenience, variety, and personalized experiences when shopping. Studies from Springer and Cushman & Wakefield indicate that consumers value the ability to shop across multiple channels—whether online, in-store, or via mobile apps—and

expect a seamless and integrated experience across these platforms. The convenience of home delivery, the ease of comparing prices and products online, and the availability of personalized recommendations are key factors that have reshaped consumer expectations. Additionally, consumer preferences vary across different demographics. Younger consumers, for example, are more likely to value digital experiences and mobile shopping, while older consumers may still prefer the tactile experience of in-store shopping but appreciate the convenience of online options. This demographic variation necessitates that retailers tailor their strategies to cater to diverse consumer needs and preferences. Customer satisfaction and loyalty are increasingly influenced by retailers' ability to provide a seamless and consistent shopping experience across all channels. Omnichannel strategies, which integrate online and offline shopping experiences, play a crucial role in building customer loyalty. According to research from JTAER and Business Insider, consumers who can easily transition between online and physical stores—whether through services like buy online, pick up in-store (BOPIS), or the availability of personalized offers across channels—tend to have higher satisfaction levels. These seamless experiences contribute to greater customer retention and loyalty, as consumers are more likely to return to retailers that meet their expectations for convenience, personalization, and service quality.

Successes & Failures

There are several case studies of brick-and-mortar retailers that have successfully adapted to the rise of e-commerce by integrating digital strategies into their operations. For instance, companies like Walmart and Target have effectively used their physical stores as hubs for online order fulfillment, offering customers options like curbside pickup and same-day delivery. These adaptations have allowed them to compete with e-commerce giants and maintain a strong market presence. The lessons learned from these successful adaptations emphasize the importance of flexibility, technological investment, and a customer-centric approach. Retailers that have embraced these strategies have not only survived but thrived in an increasingly digital landscape. However, not all brick-and-mortar retailers have successfully navigated the challenges posed by e-commerce. Some have struggled or failed to adapt, leading to store closures, financial losses, and even bankruptcy. These challenges often stem from an inability to innovate, resistance to change, or a failure to understand the shifting needs of consumers. In some cases, retailers have been slow to adopt digital tools, leading to inefficiencies and a disconnect between online and offline operations. Analysis of these failures reveals that a lack of investment in technology, poor strategic planning, and an underestimation of the impact of e-commerce are common factors that contribute to a retailer's decline. Understanding these challenges and their underlying causes is essential for other retailers looking to avoid similar pitfalls and successfully transition into the digital age.

To enhance their operational strategies in the face of e-commerce's continued growth, brick-and-mortar retailers must focus on flexibility, innovation, and a consumer-centric approach. McKinsey & Company suggests that retailers should prioritize the integration of omnichannel strategies, ensuring a seamless experience across online and physical platforms.

This includes investing in advanced digital tools, such as AI-driven personalization and real-time inventory management systems, to meet the evolving expectations of consumers. Flexibility is crucial, as retailers must be prepared to adapt quickly to changes in consumer behavior and market conditions. This could involve experimenting with new store formats, such as smaller urban locations or experiential retail spaces, that align more closely with modern shopping habits. Innovation is equally important, as it enables retailers to differentiate themselves in a crowded marketplace. Incorporating emerging technologies, like augmented reality (AR) and virtual reality (VR), can enhance the in-store experience and attract tech-savvy consumers. Additionally, retailers must maintain a strong focus on the consumer, using data analytics to better understand preferences and deliver tailored experiences that foster loyalty. Future research in the retail landscape should explore the long-term implications of e-commerce on brick-and-mortar stores, particularly in areas such as employment, urban planning, and consumer behavior. Cushman & Wakefield emphasize the importance of continuous adaptation and the need for retailers to stay ahead of emerging trends, such as the increasing importance of sustainability in consumer decision-making. Research could also examine the impact of technological advancements, such as AI and automation, on retail operations and their potential to further blur the lines between online and offline shopping. Additionally, studies on the psychological effects of digital versus physical shopping experiences could provide valuable insights into how retailers can create more engaging and satisfying customer journeys. As the retail industry continues to evolve, ongoing research will be essential to understand these shifts and guide retailers in making informed strategic decisions.

Conclusion

The rise of e-commerce has fundamentally altered the retail landscape, forcing brick-and-mortar stores to rethink their operational strategies and adapt to new consumer demands. Key findings from this paper highlight the importance of omnichannel integration, technological innovation, and a consumer-centric approach in maintaining competitiveness in the digital age. The implications for the retail industry are significant, as those who successfully embrace these changes are likely to thrive, while those who fail to adapt may struggle to survive. Looking ahead, the future of retail will be shaped by continuous innovation and a deep understanding of consumer trends, with ongoing research playing a crucial role in guiding the industry through this transformation. Ultimately, the ability to adapt to the ever-changing retail environment will determine the success of brick-and-mortar retailers in the years to come.

Works Cited

- Bihari, Lal, and Soni. "THE IMPACT of E-COMMERCE on TRADITIONAL BRICK-AND-MORTAR RETAIL BUSINESSES: A COMPARATIVE STUDY." *International Journal of Education, Modern Management, Applied Science & Social Science (IJEMMASSS)*, vol. 882, no. 04, 2023, pp. 109–16, www.inspirajournals.com/uploads/Issues/732408148.pdf.
- Chava, Sudheer, et al. "Creative Destruction? Impact of E-Commerce on the Retail Sector." *National Bureau of Economic Research*, 1 May 2022, www.nber.org/papers/w30077.
- "Bending the Cost Curve and Improving Retail Store Efficiency | McKinsey." *Www.mckinsey.com*, www.mckinsey.com/industries/retail/our-insights/bending-the-cost-curve-in-brick-and-mortar-retail.
- Boateng, Richard, et al. "E-Commerce and Socio-Economic Development: Conceptualizing the Link." *Internet Research*, vol. 18, no. 5, Oct. 2008, pp. 562–94, <https://doi.org/10.1108/10662240810912783>.
- Chandrasekar, Arjun. "E-Commerce: How It Affects Brick-And-Mortar - Finatic." 21 Feb. 2022, finatic.org/2022/02/20/e-commerce-how-it-affects-brick-and-mortar/.
- Cheba, Katarzyna, et al. "Impact of External Factors on E-Commerce Market in Cities and Its Implications on Environment." *Sustainable Cities and Society*, vol. 72, Sept. 2021, p. 103032, <https://doi.org/10.1016/j.scs.2021.103032>.
- Datta, Pratim. "A Preliminary Study of Ecommerce Adoption in Developing Countries." *Information Systems Journal*, vol. 21, no. 1, Dec. 2010, pp. 3–32, <https://doi.org/10.1111/j.1365-2575.2009.00344.x>.
- Domadenik, Polona, et al. *SHAPING the FUTURE: OPPORTUNITIES and CHALLENGES of E-COMMERCE*. 2018, www.ef.uni-lj.si/media/document_files/enote/IMB/vstavek_PKP_1.pdf#page=211.
- Laudon, Kenneth C., and Carol Guercio Traver. *"E-Commerce 2019 : Business, Technology, Society."* Thuvienso.hoasen.edu.vn, Pearson, 2020.
- Lim, Kai H., et al. "Is ECommerce Boundary-Less? Effects of Individualism–Collectivism and Uncertainty Avoidance on Internet Shopping." *Journal of International Business Studies*, vol. 35, no. 6, Sept. 2004, pp. 545–59, <https://doi.org/10.1057/palgrave.jibs.8400104>.
- Mohammed, Zareef A., and Gurvirender P. Tejay. "Examining Privacy Concerns and Ecommerce Adoption in Developing Countries: The Impact of Culture in Shaping Individuals' Perceptions toward Technology." *Computers & Security*, vol. 67, June 2017, pp. 254–65, <https://doi.org/10.1016/j.cose.2017.03.001>.
- Morrison, David E., and Julie Firmstone. "The Social Function of Trust and Implications for E-Commerce." *International Journal of Advertising*, vol. 19, no. 5, Jan. 2000, pp. 599–623, <https://doi.org/10.1080/02650487.2000.11104826>. Accessed 7 May 2020.
- Nair, Kiran. "Impact of E-Commerce on Global Business and Opportunities—a Conceptual Study." *Adu.ac.ae*, *International Journal of Advanced Engineering and Management*

- Research, 2017, repository-api.adu.ac.ae/items/4f341827-f9b2-43dc-b1b0-ba26a9d43ec9. Accessed 31 Aug. 2024.
- Sagi, John, et al. "ICT and Business in the New Economy." *Journal of Global Information Management*, vol. 12, no. 3, July 2004, pp. 44–64, <https://doi.org/10.4018/jgim.2004070103>.
- "Social Implications of Electronic Commerce." *Journal of Social Sciences*, vol. 21, no. 2, Nov. 2009, pp. 91–97, <https://doi.org/10.1080/09718923.2009.11892757>.
- Steward, S., et al. "The ECommerce Revolution." *BT Technology Journal*, vol. 17, no. 3, 1999, pp. 124–32, <https://doi.org/10.1023/a:1009640823618>.
- Tokar, Travis, et al. "A Guide to the Seen Costs and Unseen Benefits of E-Commerce." *Business Horizons*, vol. 64, no. 3, Feb. 2021, <https://doi.org/10.1016/j.bushor.2021.01.002>.

Market Analysis of Alibaba Group By Bryan Li

Basic Info

Alibaba is deeply undervalued and proves to be appealing for long-term investors because the stock is trading at a lower valuation than its expected earnings. There are numerous reasons for Alibaba's long-term expected growth including Alibaba's ongoing share repurchase program. Alibaba's plan to repurchase 25 billion dollars of their shares displays their confidence in the long-term investment of their stock as well as providing a cushion for possible future volatility or uncertainties.

Business Description

Alibaba Group is a Chinese-owned marketplace that sells products in bulk to wholesalers. The company produces money through seven different sources of income: China commerce, International commerce, Local Consumer Services, Cainiao(Alibaba's Logistics arm), Cloud, Digital Media and Entertainment, Innovation Initiatives, and others. The Alibaba group not only generates cash flow from their online wholesale marketplace but also generates money through their retail marketplace, logistic services, and a variety of other sources such as delivery of fresh goods. The Alibaba Group was founded on April 4, 1999, in Hangzhou China by Jack Ma. In Alibaba's industry, they are one of the market's strongest competitors with the highest market cap of 185.09 Billion USD. However, Alibaba's main source of revenue, international online retailing, sells products of lower quality for a significantly lower price. Because of this many individuals and businesses use Alibaba as a supplier to then resell the same products for more money. Another huge future source of revenue for Alibaba is Alibaba Cloud. Alibaba Cloud has consistently been ranked as one of the top cloud services globally for some time now. In addition, Alibaba Cloud helps diversify revenue and keeps Alibaba ahead in technology. By using resources like AI, the cloud makes Alibaba a leader in cutting-edge tech. Alibaba plays a critical role in all of this movement being one of the largest online retailers whose model is built around selling lower quality but cheaper items on the market. Alibaba's team includes the CEO, Eddie Wu, the Chairman, Joe Tsai, and the CFO, Toby Xu. Eddie Wu has been the CEO of Alibaba Group since September 23, 2023. Eddie graduated from the College of Information Engineering of Zhejiang University of Technology in 1996. Shortly after his graduation, he became Alibaba's technology director in 1999 becoming one of the co-founders at the same time. He then served as chief technology officer in 2004 at Alipay. In 2005, he became the business director of Alibaba's monetization platform and was promoted to its general manager in 2007. Later in 2008, he became the Chief Technology Officer of Taobao until he took on the role of the head of Alibaba Group search. Subsequently, Eddie was a non-executive director of Alibaba Health Information Technology Limited from 2015 to 2021 and the Chairman of Alibaba Health from 2015 to 2020. Finally, Eddie takes Daniel Zhang's spot as CEO, in 2023. Next, Jose Tsai who joined Alibaba in 1999 which makes him a co-founder like Eddie Wu. Unlike Eddie, Joe graduated from an American school, Yale, receiving his bachelor's degree in Economics and East

Asian Studies. Then, getting a juris doctor degree from Yale Law School. Before starting at Alibaba, Joe entered a couple of industries including being an associate attorney from 1990-1993 and a private equity investor from 1995-1999. Once he joined the Alibaba team he was Chief Financial Officer until 2013, Executive Vice Chairman until 2023, and recently became the Chairman in 2023. Finally, Toby Xu who has a much shorter history in the Alibaba group. He graduated from Fudan University in Shanghai, China, in 1996 with a bachelor's degree in physics. He then became a partner at PricewaterhouseCoopers for 11 years which he joined right after college until he joined Alibaba in 2018. From then until 2022 he became the deputy financial officer until he finally became the CFO in 2022. As shown, their current heads all seem to be pretty recently incorporated into their team. Jack Ma personally emphasized that these company changes aim to make Alibaba more customer-focused. These changes to the staff are sure to impact the company's future growth and how it'll develop as a whole. (Talk about the heads like the CEO, and CFO and their backgrounds, recent changes on their team, recent news like acquisitions or changes that impact their performance(developments), organizational structure of Alibaba/Alibaba Group)).

Industry Overview

Alibaba's primary source of revenue is the e-commerce market. The e-commerce market is an abundant industry that millions of people use every day. However, Alibaba does have large competitors such as PDD Holdings Inc., MercadoLibre Inc., Jd.com, and Amazon who are all primarily focused on direct-to-consumer revenue. As of 2024, global e-commerce was approximately 6.3 trillion dollars which is an 8.8% increase from the market value in 2023. However, Alibaba had an e-commerce market cap of 93.27 billion dollars in 2024, a 6.05% decrease from its market value in 2023. For a broader perspective, JD.com, one of Alibaba's biggest competitors, had a 45.46B market cap in 2023 and today they have a market cap of 64.62B marking a 43.89% increase in the last year. So, Alibaba has been underperforming compared to the rest of the market and its competitors. As shown, Alibaba has a 14.85% lower growth rate than the average e-commerce growth rate. Displaying that Alibaba is shrinking from its original massive market cap. Alibaba is losing its dominance over the e-commerce market and is instead being pretty spread out. One major advantage Alibaba has over its other e-commerce competitors is the lower prices for goods due to the manufacturing in China. In China, manufacturing and labor costs are substantially lower than other regions, therefore Alibaba can sell and deliver its products at a much lower price with a typically higher margin. Another advantage Alibaba has over its competitors is its international market. They have an international market consisting of over 200 countries worldwide making it easy for them to expand their presence overseas. Lastly, Alibaba is competitively positioned to take advantage of existing trends in increasing digital sales over companies such as Walmart. Overall their technological infrastructure is stronger than many of their competitors giving them the advantage in the upcoming years of online retail purchases. However, Alibaba has some key disadvantages too that drive buyers to have less trust in Alibaba in contrast to other big e-commerce companies.

One major disadvantage is the quality of the goods Alibaba is providing. Although Alibaba has many suppliers giving them more products and cheaper costs, some manufacturers engage in unethical business practices lowering the quality of the products. Such occasions can also lower the reputation of the company driving away potential customers. Another major disadvantage that Alibaba has is their dependence on the Chinese Market. Due to this, China's economy has a large effect on the success of Alibaba as a company. growing rapidly could be a huge advantage for Alibaba as a company. Alibaba doesn't have as much regional diversification as the other big players in the e-commerce space so investing could be riskier. However, Alibaba has invested a lot in AI, causing their cloud service to be top-notch compared to their biggest competitors. In short, AI should improve customer interactions, cut costs, and grow its cloud business, giving it a huge advantage over its competitors.

Investment Summary

Ahan Vashi wrote on Seeking Alpha that he believes that even though Alibaba is actively experiencing a little unexpected volatility, he thinks people should “view these spending sprees as necessary short-term pain for long-term gain.” Vashi is saying that right now, Alibaba is fighting off its obstacles and proving its stability while doing so. And, if you get a piece of it now, you'll make lots of money in the long term. He also says that long-term investors that are willing to put capital into Alibaba, will make an approximately 44% return on their investment in 5 years. Vashi makes a good point about Alibaba's current instability. But, if Alibaba can really get the cloud going and diversify their company well, they will be a strong investment in the long run. Next, Astrada Advisors on Seeking Alpha argued that PDD is a better investment than Alibaba because of its faster growth profile and its larger global growth potential. Astrada Advisors also claims that Alibaba is a value trap due to macro, structural, and competitive challenges. But, Alibaba proves to have much larger growth potential for many reasons including their AI technology being far ahead of PDD. If you're looking at long-term growth, AI is sure to be an essential tool for the e-commerce market, and with Alibaba's huge head start, they'll be able to seize the advantage over companies like PDD. Next, Alibaba has the benefits of economies of scale and is large enough to keep its competitive edge even though PDD has a faster growth profile. Furthermore, Alibaba isn't a value trap due to its diversified revenue base, strong leadership in key sectors, and ability to innovate. Alibaba has also proven its strong debt management and cash position proving its fundamentals are strong. Looking at Alibaba's quarterly statement, Alibaba has many diverse sources of revenue proving them to be an extremely stable company. For example, they are in international digital commerce, local services, smart logistics networks, entertainment, fresh food, as well as many other industries. Currently, they are still working to expand their AliCloud even more while making a bolstering 8 billion USD in the first half of the year. Furthermore, Alibaba “[aims] to be the leading public cloud infrastructure and platform technology provider in China”, depicting substantial growth in the industry actively making around a quarter of their revenue.

Valuation

Income Statement: Selected Financials



(in RMB Mn, except percentages)	Three months ended June 30,			
	2023	2024	YoY%	Change YoY
Income from operations	42,490	35,989	(15)%	(6,501)
Interest and investment income, net	(5,898)	(1,478)	(75)%	4,420
Interest expense	(1,784)	(2,188)	23%	(404)
Other income, net	1,364	257	(81)%	(1,107)
Income before income tax and share of results of equity method investees	36,172	32,580	(10)%	(3,592)
Income tax expenses	(6,022)	(10,063)	67%	(4,041)
Share of results of equity method investees	2,850	1,505	(47)%	(1,345)
Net income	33,000	24,022	(27)%	(8,978)
Net income attributable to ordinary shareholders	34,332	24,269	(29)%	(10,063)
Adjustments to reconcile net income to non-GAAP net income:				
Non-cash share-based compensation expense	(1,629)	4,109	N/A	5,738
Amortization and impairment of intangible assets	2,479	1,792	(28)%	(687)
Provision for the shareholder class action lawsuits	-	3,145	N/A	3,145
Loss on deemed disposals/disposals/revaluation of investments	9,038	4,581	(49)%	(4,457)
Impairment of goodwill and investments, and others	4,269	4,311	1%	42
Tax effects ⁽¹⁾	(2,235)	(1,269)	(43)%	966
Non-GAAP net income	44,922	40,691	(9)%	(4,231)
Non-GAAP net income attributable to ordinary shareholders	44,803	40,314	(10)%	(4,489)

Note:

(1) Tax effects primarily comprise tax effects relating to non-cash share-based compensation expense, amortization and impairment of intangible assets and certain gains and losses from investments, and others.

5

Alibaba's income from operations decreased by 15% however, Alibaba's net income decreased a staggering 27% from 2023. These decreases are due to the increase in non-cash share-based compensation expenses. However, Alibaba continues to function in the challenging operating environment of China. Alibaba proves to still be a strong and stable stock for long-term investors to buy in. In the market, investors are evaluating.

	Year ended March 31,		
	2022	2023	2024
	RMB	RMB	RMB
	(in millions, except per share data)		
Earnings per share			
Numerator:			
Net income attributable to ordinary shareholders for computing net income per ordinary share – basic	61,959	72,509	79,741
Dilution effect on earnings arising from share-based awards operated by equity method investees and subsidiaries	(37)	(38)	(228)
Net income attributable to ordinary shareholders for computing net income per ordinary share – diluted	61,922	72,471	79,513
Shares (denominator):			
Weighted average number of shares used in calculating net income per ordinary share – basic (million shares)	21,558	20,980	20,182
Adjustments for dilutive RSUs and share options (million shares)	229	134	177
Weighted average number of shares used in calculating net income per ordinary share – diluted (million shares)	21,787	21,114	20,359
Net income per ordinary share – basic (RMB)	2.87	3.46	3.95
Net income per ordinary share – diluted (RMB)	2.84	3.43	3.91
Earnings per ADS			
Net income per ADS – basic (RMB)	22.99	27.65	31.61
Net income per ADS – diluted (RMB)	22.74	27.46	31.24

Alibaba had an EPS(earnings per share) of \$3.92 as of October 29, 2024, a phenomenal increase from the first quarter's \$1.05 EPS. Therefore, Alibaba's current EPS leads me to believe that Alibaba will have huge profits this coming year. Furthermore, Alibaba has a current P/E ratio of 25.47 which shows the stability and strength of the company. With that being said, the ratio proves that Alibaba is in good financial health proving to be a solid investment. However, Alibaba's current free cash flow margin is down 40% from last year depicting that the company has a lot less cash on hand to use. On top of that, their free cash flow margin going down so substantially proves the struggles Alibaba is going through financially. However, Alibaba has been investing a lot into their AI and Cloud for long-term improvements in profitability and cash flow increases which gives me confidence that Alibaba is a strong long-term investment.

Cash Flow

First, this paper addresses Alibaba's cash flow from financing activities. Alibaba's RMB of 88,745 million in share buybacks demonstrates its confidence in long-term growth and undervalued stock prices. It also shows a clear commitment to returning value to shareholders. Next, Alibaba demonstrates its strong financial stability by ending FY2024 with a cash position of RMB 286,424 million despite its RMB 108,244 million in net cash flows from financing activities. They prove their ability to fund future investments, repay debt, and sustain operations. Lastly, Alibaba received RMB 20,570 million in loans from the bank while repaying large debts. This shows that Alibaba has thoughtful and wise debt management, further signaling confidence in long-term growth for the company.

Next, this paper talks about Alibaba's cash flow from investing activities. First, Alibaba has shifted its investment strategy as its net cash flow for investing strategies has decreased from RMB 135,506 million in FY2023 to RMB 21,842 million in FY2024. This could potentially mean their focus on optimizing capital allocation. Second, Alibaba has had a large increase in short-term investments, suggesting their more conservative approach rather than riskier long-term high-risk investments. Alibaba reported a net increase of RMB 71,426 million in short-term investments in FY2024. Lastly, Alibaba continues its investment in Equity Securities and equity method investees, despite reduced cash outflow. By doing so, Alibaba showcases its focus on strategic partnerships and growth opportunities.

Lastly, this paper reviews Alibaba's cash flow from operating activities. To start, Alibaba has strong growth in its net income. Their net income grew from RMB 65,573 million in FY2023 to RMB 71,332 million in FY2024, portraying a healthy financial performance. Next, Alibaba has a lot of depreciation and impairment for their property. They have had a total of RMB 26,640 million in depreciation of equipment and property showing ongoing investments and asset usage. Lastly, Alibaba's net cash from operating activities was RMB 182,593 million in FY2024. Showing that despite fluctuations in expenses, Alibaba is still able to generate cash from its core business.

Financial Analysis

Alibaba's strong and stable cash flow and consistent profitability and growth shows great potential for the future of the company in a long-term scope. They also have adapted well in challenging macroeconomic environments and show disciplined capital management. Next, with over 1 billion active Chinese consumers, they provide a reliable stream of transaction feeds and advertising revenue. But, not only are they dominant in the e-commerce market, but they are also commanding a huge share of China's cloud market with AliCloud. With these 2 major roles, Alibaba displays their leadership in the markets. In the year, Alibaba had approximately 138 billion USD in revenue displaying a strong 5.14% year-over-year growth. Additionally, the company had a great gross profit of 38.2% of revenue with a operating margin of 13.71%. Next, Alibaba has a P/E ratio of 16.2x, suggesting that Alibaba is undervalued relative to its earnings. On top of that, Alibaba has a FCF margin of a robust 10.76% demonstrating their ability to

generate cash from operations as well as having enough cash in hand to spend on strategies. The company also expanded their international retail operations, with a 35% increase in revenue to around 3.7 billion USD. Lastly, the company continues to make strategic investments, investing in technology like cloud computing and infrastructure to enhance their competitive edge; Overall, they have some risks they need to overcome, but if they can diversify revenues through cloud computing successfully, it'll be a big step forward for sustaining long-term growth.

Investment Risks

First, Macroeconomic uncertainty is a risk when investing in Alibaba. Since Alibaba operates in many diverse industries, the company becomes vulnerable to global economic conditions. Therefore things like global inflation or slower GDP growth could directly impact Alibaba's core e-commerce business. Secondly, Competitive pressure in e-commerce and cloud computing affects Alibaba as well. Alibaba has lots of major competition in both their own region and globally like JD.com and PDD. Because of this competition, Alibaba is forced to continuously invest more money into different industries of their business to keep up. Next, Alibaba faces Geopolitical tensions due to being a Chinese company. They are tied to China and therefore can be limited from expanding internationally from global tensions. Additionally, they could face market restrictions and limitations on market access. On top of that, being a major player in China, an economic slowdown in China could cut their revenues substantially. Another pressure the business must face is being profitable. Since Alibaba is investing in many new actions like international growth and AliCloud; they could face lower margins making it harder to sustain earnings to grow and flourish as a company. AliCloud in specific can encounter many challenges, for example, slower than anticipated adoption of Cloud in emerging markets. Therefore, Cloud could hinder the growth of the company. Lastly, Alibaba's engagement in buybacks may lead to assurance for investors in the short-term. However, it could limit the company's financial flexibility for other strategic investments. Over time, excessive buybacks could also lead to criticism regarding the effective use of capital.

Conclusion

Overall, Alibaba proves to be a stable company that's profitable. Alibaba also has a lot of faith in their own company to grow and expand by buying back their own stock which reassures investors with long-term growth. Alibaba has demonstrated their AI edge over their biggest competitors. They intend to continue diverting their investment into Cloud, one of their biggest edges, to win over the e-commerce market in the long run. Although Alibaba has some risks like some of the biggest ones begging geopolitical tensions and high share buybacks. Many analyzers believe that Alibaba will be able to overcome these risks and become the leading company in e-commerce with the help of Cloud. Therefore you should buy Alibaba before it's too late.

Works Cited

<https://www.precedenceresearch.com/e-commerce-market>
<https://hbr.org/2018/09/alibaba-and-the-future-of-business>
<https://blog.osum.com/alibaba-group-swot-analysis/#:~:text=Weaknesses%20of%20Alibaba%20Group,-While%20Alibaba%20Group&text=These%20weaknesses%20include%3A,Chinese%20economy%20and%20regulatory%20environment.>
<https://www.forbes.com/sites/greatspeculations/2024/07/17/whats-happening-with-alibaba-stock/#:~:text=The%20regulatory%20troubles%20that%20Alibaba,Alibaba's%20valuation%20is%20also%20compelling.>
<https://pitchgrade.com/companies/alibaba#:~:text=Alibaba's%20competitive%20advantage%20lies%20in,development%2C%20marketing%2C%20and%20logistics.>
<https://www.linkedin.com/pulse/exploring-pros-cons-leveraging-alibaba-your-sourcing-needs-popescu-mso4f#:~:text=Disadvantages%3A,of%20quality%20control%20and%20reputation.>
<https://pitchgrade.com/companies/alibaba>
<https://www.marketbeat.com/stocks/NYSE/BABA/earnings/>
<https://finance.yahoo.com/quote/BABA/>
<https://www.alibabagroup.com/en-US/document-1761110461375315968>
<https://stockanalysis.com/stocks/baba/>
<https://www.wsj.com/market-data/quotes/BABA/research-ratings>

Sociocultural Norms Shaping Sports Participation Among South Asian & Arab Female Athletes

By Devika Hajarnavis

Abstract

The struggles of female athletes in South Asia and the Gulf region are often overshadowed in popular media by challenges faced by women athletes in the West. This paper sheds light on the societal, cultural, and religious barriers that hinder the participation and potential success of female athletes in South Asia and the Gulf region. It also highlights the resilience and perseverance demonstrated by these female athletes in order to push back against systemic social and cultural norms.

This paper deliberately avoids adopting a “White Savior Complex” perspective when examining the regionally specific issues faced by female athletes. It emphasizes the importance of culturally sensitive solutions that aim to support these athletes without tearing them away from their traditional or religious values.

With a primary focus on India and Saudi Arabia, this paper advocates for a more inclusive approach to sports that respects and acknowledges the different cultural backgrounds that female athletes come from, and identifies the problems they face, while giving examples of successful female-athletes. Advocating to amplify the voices of female athletes, this paper gives solutions for a more inclusive and supportive environment in sports.

Introduction

Athletes are often referred to as “fighters” signifying their competitive spirit in the playing arena and their resilience, tenacity and determination in the pursuit of excellence in their sport. However, in certain cultures, women athletes face a very different fight: in their homes, communities and cultures, where they face uphill struggles for the right to play their sport.

In 2023, Indian female wrestlers faced their toughest fight – not on the wrestling mat, but against their governing body. On the streets of Delhi, three of India’s most renowned wrestlers—including two Olympic medalists—were forcefully dragged by the police as they protested against the President of the Wrestling Federation of India (WFI), Brij Bhushan Singh, who was accused of mentally harassing and sexually assaulting several female athletes (Top Indian female wrestler quits in protest over new president of wrestling body, 2023). Highly decorated wrestlers (female and male) protested against him demanding action to be taken. In response to these protests, Brij Bhushan Singh was replaced in the WFI by Sanjay Singh, one of his closest aides (Top Indian female wrestler quits in protest over new president of wrestling body, 2023). This move caused significant uproar – with wrestlers protesting on the roads, fighting against the political establishment and demanding an independent investigation (Vasudevan, 2023). This example of WFI shows how the safety of female athletes has been undermined by governing bodies of sports associations.

For girls in the West, getting a driver’s license while still in their teens is a rite of passage. However, until June 2018, women were forbidden from driving in Saudi Arabia

(Specia, 2019). Against the backdrop of that restriction, you have the example of Reema Juffali, a Saudi national who was introduced to motorsport when she moved to the US and fell in love with it (Making history from the cockpit: Juffali on becoming Saudi Arabia's first female racing driver., 2019). Ultimately, Juffali became the first woman from Saudi Arabia to hold a racing license, although by her own admission, her late start to the game impacted her growth curve (<https://news.northeastern.edu/2019/08/06/making-history-from-the-cockpit-reema-juffali-on-becoming-saudi-arabias-first-female-racing-driver/>, 2019). Moreover, she did wonder what others in her culture would think as she knew her path was “something that's foreign to them, something that hasn't been around in their society or around them (Roper, 2023).”

Established socio-cultural values and religious beliefs in South-Asian and Gulf nations have had an overbearing and detrimental impact on the participation and performance of women athletes. For the female athletes from these nations who have attained success and global recognition, their achievements come from breaking through the shackles of systemic social and cultural norms. Deeply held socio-cultural or religious beliefs within certain communities have traditionally demanded that women sacrifice sporting dreams and, in some cases, even budding interests. There is a gnawing need to develop a more inclusive approach to sports that amplifies the voices and encourages the dreams of female athletes, while recognizing that socio-cultural beliefs in regions will imply that the path to inclusivity may not be smooth and might need milestones along the way.

This paper intentionally has a different scope in comparison to several other papers about female athletes: highlighting issues faced by women athletes in the Gulf and South Asia, without adopting a Western frame of reference. This does not mean that female athletes in Western regions do not face gender-based issues; however, several reports have focused on those issues and identified solutions, which is promising. In contrast, the relative silence on issues faced by female sports aspirants and athletes in South Asia and the Gulf is deafening. It is crucial to be culturally sensitive to the issues and potential solutions one speaks about in different regions due to different contexts. This paper aims to examine the situation with that context in mind and not with a “White Savior Complex.”

When examining South Asia, India will be the primary focus. As the most populous nation in the world, India dominates the South Asian sports world (Mashal, 2024). When examining the Gulf, the focus will be on Saudi Arabia as the country represents a desired transformation from the extreme gender barriers in sports that are prevalent in the region to increased women’s participation with recent changes in rules.

Despite significant cultural, societal, and religious barriers, female athletes in South Asia and the Gulf region demonstrate remarkable resilience and determination; this paper explores the impact of obstacles on their careers, argues against forcing a choice between athletic aspirations and traditional values, and advocates for culturally sensitive solutions to support and empower these athletes. This paper will delve into five different issues that female athletes and sports women face, specific to the Gulf and South Asia. This includes issues of (i) gender-based stereotyping, (ii) limited early exposure to physical development and exercise, (iii) limited

sporting infrastructure and the lack of coaching resources, (iv) limited amount of awareness and recognition of female athletes, and (v) concerns about the safety of women. I will also explore potential solutions for these problems that can accommodate female athletes from diverse sociocultural backgrounds, allowing them to participate in sports without compromising their cultural values.

Issues faced by female sportspersons and sports aspirants

1. Gender-Based stereotyping

As a young girl, Mithali Raj accompanied her brother to his cricket training and did her homework at the boundary lines. When bored, she picked up a bat and started knocking the ball around. Noticed by a coach for her technique and supported by her parents, she received training, and her interest in the game turned into a passion. Pursuing that passion, however, wasn't easy as her grandparents discouraged her from pursuing a "boys' sport (Patwardhan)." In Jodhpur, Rajasthan, during the 1990s, socio-cultural norms meant that girls were discouraged from playing sports and directed towards household responsibilities (Patwardhan, 2022). Raj's journey was challenging: her practice sessions were solitary, a symptom of the prevalent norms and the lack of sporting infrastructure for girls.

Gender-based societal and cultural expectations begin at a young age when children receive early conditioning about their roles. This is true in general and not only in South Asia. For instance, stereotypically, young boys receive balls or cars as presents from their family, while girls receive dolls —feeding into the bias that boys are natural at sports while girls are natural at caregiving. Research by various sources throughout the years has shown that parents and adults tend to have different behaviors towards children, even newborns, according to the sex of the child (Hargreaves, 1994). Girls tend to be handled more gently and are supervised closely with less freedom, while boys are handled more aggressively and have more physical freedom. In families, gender stereotyping can be perceived as normal. Children quickly understand these set ideas and "roles (Hargreaves, 1994)" they must follow based on their assigned sex. Margaret MacNeill argued that "children learn that certain activities are more appropriate for girls and others are more suitable for boys", without ever being explicitly taught or told that (Hargreaves, 1994). In her 2002 paper "Sporting Females," Jennifer Hargreaves provides examples of this phenomenon in the way that sports are partitioned between boys and girls in England (soccer vs. gymnastics) and Canada (ice hockey vs. figure skating). Boys are usually more encouraged to participate in sports, and even when girls are encouraged, they typically are in smaller numbers and nudged towards sports characterized as "aesthetic and expressive" such as gymnastics or ice skating, compared to "aggressive and tough" sports like soccer or hockey (Hargreaves, 1994).

In March 2022, the Pew Research Center (How Indians View Gender Roles in Families and Society." Pew Research Center., 2022) surveyed nearly 30,000 Indians (gender breakdown of the group unavailable) to explore various aspects of gender roles within families and societies. The findings of the study were surprising and presented somewhat contradictory results. While

most respondents believed that men and women are equally capable as leaders, with 14% expressing that women are even better leaders than men, there was also a strong consensus within those surveyed regarding the notion of women needing to obey their husbands, with 64% in agreement and 24% mainly in agreement. This dangerously misogynistic belief, which can be harshly visible in rural parts of India (and a prevalent bias in urban settings), is a stark reflection of the social and cultural values. For instance, if a woman's husband disapproves of her participation in sports, Indian socio-cultural values would normalize the woman's prioritization of her husband's approval over her desires and aspirations.

2. Limited early exposure to physical development and exercise

In Saudi Arabia, some religious leaders believed that girls should not go to school, and even basic education for them was opposed (Zambarakji, 2012). Fortunately, since 2009, it has been made mandatory for all children in Saudi Arabia to go to school until they are 16, including girls (Education in Saudi Arabia." World Education News + Reviews, 2020). Nonetheless, the freedom of girls in Saudi Arabia remains highly curtailed. For instance, women need a male relative's permission for life choices such as marriage, getting a job, or obtaining a passport. It was only after 2017 that girls in Saudi Arabia could receive physical education lessons as part of the public-school curriculum. Until then, there was a nation-wide ban on sports for girls in public schools. This change came four years after the removal of the restriction on sports participation for girls in private schools (Saudi government sanctions sports in girls' private schools, 2013).

While many have welcomed the removal of the ban in Saudi Arabia that severely limited the ability of girls to participate in even basic fitness activities, there was no doubt that the change in mindsets would take time. The ban was indicative of the predominant conservative view that opening sports to women and girls will lead to immorality. In the words of one religious scholar, it was the "steps of the devil ('Steps of the Devil' - Denial of Women's and Girls' Rights to Sport in Saudi Arabia, 2012)." Opponents of allowing participation in sports for women and girls in Saudi Arabia put forward the "slippery slope" argument that once women start to exercise, they will shed modest clothing, spend "unnecessary" time out of the house, and have increased possibilities for mingling with men ('Steps of the Devil' - Denial of Women's and Girls' Rights to Sport in Saudi Arabia, 2012).

In India, with limited resources, the ability of girls to receive physical education training in school is highly restricted. This issue is even more pronounced in rural settings, where nearly two-thirds of the population lives. According to the 2011 Census of India, only one out of every 100 girls in rural India reaches twelfth grade. A non-profit organization, Educate Girls, discovered that 2.5 million Indian girls have never been to school in part because only half of the public schools have toilets for girls (What's the secret weapon keeping girls in school in India, 2024). The accessibility of public schools and the high dropout rates have meant that many rural Indian families have subscribed to the mindset of keeping their daughters at home and teaching them how to run a household with the goal of getting them prepared for marriage.

3. Limited sporting infrastructure for women, and a lack of coaching resources

Even after Raj was selected for the Indian women's cricket team in 1999, she faced numerous obstacles on her journey to becoming a better player. While the men's team received all necessary support with the backing of the governing board's coffers, the women's team struggled for facilities, coaching, and even nutrition, as expectations, viewership and resultant sponsorships for the women's team were limited ('I hope I see the day when people acknowledge men's and women's cricket equally., 2020). With meager resources, the choice for many of the national team women players was to support themselves with part-time government jobs. Raj herself took on a government job, thereby reducing her training at a time when she needed to focus exclusively on her sport.

Mirabhai Chalu's potential was first spotted by her father, who said "she used to lift a bundle of fire logs easily at the age of 12 as compared to her elder brother (Inspiring story of Mirabhai Chalu, the Silver Medalist from India in Tokyo Olympics 2020, 2020)." Chalu (who hails from a tribal area in India) went on to win a silver medal in weightlifting at the 2020 Tokyo Olympics in Japan, but her father's comment highlights how her informal talent spotting and training began.

For years as an economically developing nation, India did not invest adequately in sporting development and infrastructure (The Economic Transcript. "Inadequate Infrastructure for Athletes in India: A Roadblock to Success, 2024). That has changed over the last decade with programs like *Khelo India Youth Games* (Khelo India: All you need to know about the Youth Games, University Games and the Winter Games, 2024), but the patriarchal approach has meant that investment in sports facilities for women especially in rural areas has remained low. While the women's cricket team now has resources and recognition —thanks to the growing popularity of the sport in India and the increased sponsorship for women's cricket worldwide— Raj's journey illustrates the real struggles that she and her teammates had to go through at a time when the men's team was well-funded.

Saudi Arabia's challenge hasn't been the lack of resources for stadium infrastructure but the cultural restrictions that didn't allow women to enter them in the first place. It was only in 2018, that women were finally allowed to watch sporting events in the family section (Saudi Arabia allows women at football game for first time, 2018).

In 2017, the Saudi Arabian government implemented the Vision 2030 plan: a comprehensive list of goals the country wants to achieve by 2030 (Vision 2030, 2024). The goals are built upon three primary themes: a vibrant society, a thriving economy, and an ambitious nation. One of the key tenets behind Vision 2030 is an aim to diversify the economy while promoting social reforms and women's empowerment. This aim has resulted in a significant transformation in its stated outward approach to women's participation in sports. The Kingdom sees the growth of the sports sector for both men and women as key to accelerating economic diversification, increasing the inflow of foreign investment and enhancing the attractiveness of Saudi tourism. As a result, the country has since been on a rapid trajectory with "announced" changes to women's sports (Vision 2030, 2024). Even a women's soccer league has been

established, although most players are from other countries. That is not surprising as with the most basic of physical education resources just opening up for girls, and the mindset still engrained in conservatism, it will take time for the women of the country to develop their talent and competence.

While the goals for Vision 2030 are admirable, changing the mindsets of people in Saudi Arabia through education and awareness continues to be essential, particularly in light of comments made by influential leaders in the country. For instance, in 2009, Sheikh Abdullah al-Manee'a, a member of the official Supreme Court of Religious Scholars in Saudi, claimed that the “movement and jumping” that sports (such as football and basketball) require may cause hymen tears for girls, which can “give the appearance that they have lost their virginity”(Zambarakji, 2012). As an offshoot of this mindset, Saudi women lived a sedentary lifestyle, resulting in 45% of women being obese, “far greater than among men” (in 2010) (Zambarakji, 2012). To fix the national obesity problem, Saudi Arabia did two things. In 2017, the Kingdom opened women-only gyms, which encouraged Saudi women to adopt a healthy lifestyle while limiting their contact with males (“Saudi women can finally hit the gym, 2017). By 2019, only 21.4% of women were obese, a vast improvement from previous years (Ministry of Health - Saudi Arabia, Obesity, H.R. Rep, 2024). Secondly, the nation focused on increasing the participation in sports from 13% to 40% of the population, and most importantly, including women in that goal.

“4 out of 64 — Where are the women coaches in Indian national teams?”(4 out of 64 — Where are the women coaches in Indian national teams, 2021). In the 2020 Tokyo Olympics, India participated in 64 events across 27 sports. The number of women coaches associated with those events was limited to only four. This was on the heels of the Rio Olympics, where India’s two medals were both won by women (PV Sindhu and Sakshi Malik).

Strong coaching is fundamental to the growth of athletes, and in India coaching in women’s sports has played a distant second fiddle to men’s sports. This issue has been exacerbated by the lack of female coaches at the grassroots level due to women not having been supported in their sporting careers. Furthermore, systemic issues play a part in the sustainability of coaches as seen in Hockey India.

Janneke Schopman, an ex-Dutch Hockey player, took over as coach of the Indian women’s hockey team in 2020, becoming the first-ever female hockey coach of an Indian team (India's women hockey team coach Janneke Schopman says India 'extremely difficult for women, 2024). She had followed Sjoerd Marine who had spoken about how his reward for good work with the women’s hockey team was getting promoted to get the men’s team job (Sjoerd Marijne quits as Indian women's hockey team coach, 2021) – and then subsequently, after bad results with the men’s team, he was removed from that job and re-appointed women's coach. On February 18, 2024, Schopman claimed she was facing “inferior treatment” in India due to her gender (Extremely difficult as a woman: Indian women's hockey coach Schopman alleges gender discrimination, 2024). Speaking against Hockey India, the governing body for the national men’s and women’s hockey teams, Schopman said, *“I come from a culture where women are respected*

and valued. I don't feel that here... I see the difference in how the men's coaches are treated and between the girls and the men's team...this country is extremely difficult as a woman." On February 23, 2024, she resigned as Head Coach (Janneke Schopman resigns as head coach of Indian women's hockey team, 2024). When interviewed by NNIS Sports, she said, *"I always thought I would get fired if I spoke the truth"*(I always thought I would get fired if I spoke the truth: Schopma, 2024)." A few days later, Elena Norman, Hockey India's first woman CEO, stepped down from her position, claiming that working at Hockey India was a "suffocating" work environment (It was becoming suffocating: Hockey India CEO Elena Norman on resignation, 2024). Norman's 13 years at Hockey India had not only helped bring the women's team into the nation's spotlight and provided equal facilities for the women's team, but it was under her leadership that the men's team had achieved a bronze medal at the Olympics ending a 41-year medal drought in the country's national sport (Hockey CEO Elena Norman's resignation is a big deal, for hockey and for women in sport., 2024). Shockingly, Norman and Schopman's disturbing allegations in light of their resignations were not investigated. Hockey India simply issued a statement refuting their allegations (It was becoming suffocating: Hockey India CEO Elena Norman on resignation, 2024) and the matter was closed.

While female sportspersons are bucking the trend and showing their skills in the competitive arena, India still has a long way to go when it comes to producing women in coaching and other support staff roles. At the grassroots level, it is an evolutionary process – only once women have spent enough time playing and training themselves do they usually consider going into coaching. The lack of awareness and opportunity means that there isn't a well-established blueprint or clear procedures to follow for women to follow if they wish to enter coaching and training.

4. Limited awareness and recognition

Could you name a female athlete, if asked? Most likely, you could. However, a surprising 50% of Indians surveyed by the BBC in 2020 could not name a single female Indian athlete (What do Indians think about women in sport, 2020). This statistic is not just surprising but also disheartening. At both the Tokyo and Paris Olympics, Indian women have won nearly half the number of medals for the country (Inspiring story of Mirabai Chanu, the Silver Medalist from India in Tokyo Olympics 2020, 2020). The same research conducted by the BBC indicated that a third of Indians believe sportswomen are not as good as sportsmen. The research indicates that India is facing a significant issue—not only is it not promoting the development of women in sports, but it is also failing to recognize its female champions.

When participants in the BBC survey were asked about sports considered 'unsuitable for women,' activities traditionally perceived as 'masculine,' such as weightlifting, wrestling, and boxing, were frequently mentioned (New BBC research shows less than 30% of Indian women play any sports, 2020). Ironically, several female athletes from India have defied stereotypes and achieved remarkable success in these very sports, winning accolades at Commonwealth Games, Asian Games and the Olympics (Navyasari. Women Entrepreneurs (blog), 2022). Santanu

Chakrabarti, Head of Audience Research, BBC World Service, said: "Our research shows attitudes towards women and women's sports in India are complex, contradictory and counterintuitive."

The disconnect in mindsets was also evident by the statistic that while over three-quarters of the people surveyed believe that participating in sports is essential, only 29% of the women surveyed said they played any sport. The researchers concluded that while some progress had been made in terms of attitudes towards women's sports thanks to the achievements of sportswomen, age-old stereotypes still prevailed. Nearly two-fifths of respondents in the BBC survey believed that women's sports were not as entertaining as men's sports. This translates to low viewership and sponsorships for women athletes creating a vicious cycle of funding constraints.

Mitali Raj holds the honor of being the first Indian cricket captain, regardless of gender, to have led their team to two prestigious World Cup finals, a testament to her trailblazing spirit. Despite that, Raj is sometimes asked to name her favorite male cricketer in interviews. Her quick retort is, "Do you ask male cricketers who their favorite female cricketer is? (Patwardhan, 2022)"

5. Concerns about the safety of girls

The importance of having enough female coaches cannot be overstated in increasing the participation of girls. For the girls and their parents, physical security and the creation of a safe space while they play and learn is of prime importance. Unfortunately, with India's spotted track record on this issue, concerns for their daughters' physical safety are a key reason for families not allowing them to pursue sports as that often requires a commute to a training site and travel to competitions. The paucity of female coaches supervising and supporting young girls in their sports journey contributes to those concerns.

Potential Solutions

This section of the paper will delve into potential solutions and actions that can be taken to enhance the participation of Arab and South Asian female athletes. The main steps that need to be taken are 1) a change in mindset, 2) an increase in the number of female coaches and a focus on infrastructure, and 3) a focus on girls' safety.

1. A mindset shift

A mindset shift is crucial because it reduces harmful sexist stereotypes in a region. These stereotypes limit the potential of female athletes and contribute to a cycle of inequality in sports. The first step to address and combat these stereotypes is to create an inclusive and supportive environment for young Arab and South Asian girls. For a mindset shift to occur, there needs to be an increase in awareness of the opportunities in sports and an increase in role models for young girls.

To increase awareness of opportunities, sports organizations and authorities need to help in creating mentorship programs and sports events targeted explicitly at young girls and

implement policies to promote gender equality in sports. Returning to Saudi Arabia, Vision 2030's ambitions have supported rapid growth in the sports sector. Female participation in sports has increased by 150% since 2015 (Female participation in sports up 150% in Saudi Arabia, 2021). This growth represents a powerful shift in mindset, which signals the potential for a progressive future. These initiatives can only occur with institutional support and a focus on women's participation. An example of a youth program that is flourishing in South Asia is *The Naandi Foundation's Game Changer Program*, which is focused on empowering rural girls in India. The program teaches both young girls and women how to stay active, and helps them become certified football coaches (How More Women Coaches Can Change the Game in India, 2022).

To increase the number of mentors and role models for younger girls, nations need to celebrate the progress of their current female athletes and sportswomen. In India, where

Bollywood movies are a massive part of the culture, more content about female athletes should be made. In previous years, movies such as "*Chak De India*," "*Saina*," "*Mary Kom*," and even "*Dangal*," did a commendable job of showing not just the hardships Indian female athletes had to go through but also their achievements on international stages, which inspired young girls to follow in their footsteps. After seeing how those top athletes broke barriers in sports such as badminton, wrestling, boxing, and shooting, parents enrolled children in classes to see if their young girls could become the next 'World Number 1' (Sharma, 2015).

2. Increase in female coaches and a focus on infrastructure

Countries in the Gulf and South Asia have not focused enough on their female athletes in the past; hence, they do not have female coaches with the same amount of experience in the sport as male coaches. Therefore, they must hire female coaches from other countries to support their athletes and provide them with the infrastructure and support to train young girls.

In both the Gulf and South Asia, infrastructure has two levels: 1) at the grassroots level and 2) at the professional level. At the grassroots level, facilities in rural areas and schools need to improve through funding. At the professional level, nations must ensure that their professional resources and training arenas are of the highest quality to support their athletes. This means ensuring athletes can access necessary equipment, nutritionists, athletic trainers, etc.

3. A focus on girls' safety

Although a lack of focus on safety of female athletes is not unique to a specific region, it is crucially important to discuss. Sadly, female athletes across regions have often suffered abuse and the fear can keep women and girls away from sports. The way to improve this, is by educating coaches and authorities on how to care and train for their female athletes while keeping them safe from the different forms of abuse (whether that is physical, sexual, mental), and making sure they are comfortable while training. Furthermore, any abuse reported by athletes must be investigated independently and swiftly with corrective actions taken to safeguard the safety of the athletes. Female athletes can be made comfortable in training arenas

by making the sporting environment inclusive, supportive and empowering. An increase in the number of women coaches and women in support staff roles would also support this.

Conclusion

The journey for female athletes in South Asia and the Gulf has not been easy. Societal, cultural, and religious barriers have hindered the participation and development of women in sports. Gender-based stereotyping, limited opportunities to participate in sports due to poor infrastructure, lack of women role-models and coaches and safety concerns have all held back women athletes. Despite those challenges some sportswomen have succeeded at the international level. However, it begs the question of how many more could have secured their sports dreams if they had received the right support. While keeping cultural sensitivities in mind, this paper has identified solutions for enhancing sporting infrastructure and accessibility at the grassroots level and changing mindsets to encourage more girls to participate and thrive in sports. With positive changes underway in women's sports in these regions, the hope is that structural changes can provide the much needed impetus and help many more women achieve their sporting dreams.

Works Cited

- "I always thought I would get fired if I spoke the truth: Schopman." Interview. Youtube.
<https://www.youtube.com/watch?app=desktop&v=wDbLwCTlhKs>.
- "I hope I see the day when people acknowledge men's and women's cricket equally." *The Cricket Monthly*.
<https://www.thecricketmonthly.com/story/1219778/-i-hope-i-see-the-day-when-people-acknowledge-men-s-and-women-s-cricket-equally>.
- Al-Khalifa, Hussa K. "We're all in this together': Perspectives from within the Gulf Cooperation Council Women's Games." *Sage Journals* 57, no. 2 (2021).
<https://doi.org/10.1177/101269022199>.
- "Female participation in sports up 150% in Saudi Arabia." *Arab News* (Riyadh, Saudi Arabia).
<https://www.arabnews.com/node/1785526/sport>.
- "New BBC research shows less than 30% of Indian women play any sports." *BBC*.
<https://www.bbc.com/mediacentre/latestnews/2020/indian-women-play-any-sports>.
- "Saudi Arabia allows women at football game for first time." *BBC*.
<https://www.bbc.com/news/world-middle-east-42668841>.
- "What do Indians think about women in sport?" *BBC*.
<https://www.bbc.com/news/world-asia-india-51701924>.
- "It was becoming suffocating: Hockey India CEO Elena Norman on resignation." *Business Standard*.
http://www.business-standard.com/sports/other-sports-news/it-wasbecoming-suffocating-hockey-india-ceo-elena-norman-on-resignation-124022700604_1.html.
- Chak De India!* Directed by Shimit Amin. 2007.
- "Why is India so bad at sport?" *CNBC*.
<https://www.cnn.com/2016/08/19/lack-of-sporting-culture-institutional-support-and-inequality-blamed-for-indias-poor-olympic-record.html>.
- "Saudi government sanctions sports in girls' private schools." *CNN*.
<https://www.cnn.com/2013/05/05/world/meast/saudi-arabia-girls-sports/index.html>.
- Dangal*. Directed by Nitesh Tiwari. 2016.
- The Economic Transcript. "Inadequate Infrastructure for Athletes in India: A Roadblock to Success." *The Economic Transcript*.
<https://tetoofficial.com/inadequate-infrastructure-for-athletes-in-india-a-roadblock-to-success/>
- "Education in Saudi Arabia." *World Education News + Reviews*.
<https://wenr.wes.org/2020/04/education-in-saudi-arabia>.
- "Extremely difficult as a woman: Indian women's hockey coach Schopman alleges gender discrimination." *ESPN*.
https://www.espn.com/field-hockey/story/_/id/39557158/janneke-schopman-indian-women-hockey-coach-alleges-gender-discrimination.

- "Hockey CEO Elena Norman's resignation is a big deal, for hockey and for women in sport."
https://www.espn.in/field-hockey/story/_/id/39609560/hockey-india-ceo-resignation-elen-a-norman-big-deal-hockey-women-sport.
- "Janneke Schopman resigns as head coach of Indian women's hockey team." *ESPN*
https://www.espn.com/field-hockey/story/_/id/39583334/janneke-schopman-resigns-head-coach-indian-women-hockey-team.
- "Sjoerd Marijne quits as Indian women's hockey team coach." *ESPN*
https://www.espn.in/olympics/fieldhockey/story/_/id/31970560/tokyo-olympics-sjoerd-marijne-quits-indian-women-hockey-team-coach.
- "Female athleticism on rise in Qatar and Middle East, accompanied by social and cultural challenges." *Northwestern University*.
<https://qatar.northwestern.edu/news/articles/2012/02-female-altheticism-gulf.html>.
- "4 out of 64 — Where are the women coaches in Indian national teams?" *The Bridge*.
<https://thebridge.in/others/4-women-coaches-india-national-teams-20002>.
- Hargreaves, Jennifer. *Sporting Females: Critical Issues in the History and Sociology of Women's Sport*. N.p., 1994.
- "How More Women Coaches Can Change the Game in India." Interview. Video. YouTube.
 Posted by The Bastion, April 9, 2022.
https://www.youtube.com/watch?v=4IHHSoS9PTM&ab_channel=TheBastion.
- "Saudi Arabia: Women Are 'Changing the Game.'" *Human Rights Watch*.
<https://www.hrw.org/news/2016/08/04/saudi-arabia-women-are-changing-game>.
- "'Steps of the Devil' - Denial of Women's and Girls' Rights to Sport in Saudi Arabia." *Human Rights Watch* (New York City, NY), February 15, 2012.
- "In Saudi Arabia, women couldn't drive cars until last year. This Saudi woman races one."
<https://cssh.northeastern.edu/in-saudi-arabia-women-couldnt-drive-cars-until-last-year-thi-s-saudi-woman-races-one/>.
- "Inspiring story of Mirabai Chanu, the Silver Medalist from India in Tokyo Olympics 2020."
Inspiration Unlimited.
<https://www.iuemag.com/n22/is/inspiring-story-of-mirabai-chanu-the-silver-medalist>.
- Making history from the cockpit: Juffali on becoming Saudi Arabia's first female racing driver*.
<https://news.northeastern.edu/2019/08/06/making-history-from-the-cockpit-reema-juffali-on-becoming-saudi-arabias-first-female-racing-driver/>.
- Maktabi, Rima, and Schams Elwazer. "Saudi women: Pampered or oppressed?" *CNN*.
<https://edition.cnn.com/2012/03/14/world/meast/saudi-women-disagree-rights/index.html>.
- Mary Kom*. Directed by Omung Kumar. 2014.
- Mashal, Mujib. "India Wins Cricket World Cup, Sealing Its Domination of the Sport." *The New York Times*.
<https://www.nytimes.com/2024/06/30/world/asia/india-wins-cricket-world-cup.html>.
- "India Wins Cricket World Cup, Sealing Its Domination of the Sport." *The New York Times*.
<https://www.nytimes.com/2024/06/30/world/asia/india-wins-cricket-world-cup.html>.

Ministry of Health - Saudi Arabia, Obesity, H.R. Rep.
<https://www.moh.gov.sa/en/awarenessplatform/ChronicDisease/Pages/Obesity.aspx>.

"India's women hockey team coach Janneke Schopman says India 'extremely difficult for women.'" *Mint*
<https://www.livemint.com/sports/news/it-is-very-difficult-for-women-to-work-in-india-11708350314151.html>.

Mogul, Rhea. "Top Indian wrestler quits after federation elects ally of controversial former chief." *CNN*.
<https://www.cnn.com/2023/12/22/india/india-sakshi-malik-quits-wrestling-intl-hnk/index.html>.

Navyasari. *Women Entrepreneurs* (blog).
<https://www.womenentrepreneursreview.com/viewpoint/experts-column/female-athletes-who-won-big-made-india-proud-in-the-common-wealth-games-2022-nwid-2240.html>.

"Saudi women can finally hit the gym." *The New Arab*.
<https://www.newarab.com/features/saudi-women-can-finally-hit-gym>.

"Saudi Arabia to Offer Physical Education Classes for Girls." *The New York Times*.
<https://www.nytimes.com/2017/07/11/world/middleeast/saudi-arabia-girls-islam-physical-education.html>.

"Khelo India: All you need to know about the Youth Games, University Games and the Winter Games."
<https://olympics.com/en/news/khelo-india-games-youth-university-school-history-winners>.

Patwardhan, Deepti. "Mithali Raj: The 'sleepy' girl who changed Indian cricket forever." *BBC*.
<https://www.bbc.com/news/world-asia-india-61741814>.

"How Indians View Gender Roles in Families and Society." *Pew Research Center*.
<https://www.pewresearch.org/religion/2022/03/02/how-indians-view-gender-roles-in-families-and-society/>.

"Saudi Arabia to introduce physical education for schoolgirls." *Reuters*
<https://www.reuters.com/article/world/saudi-arabia-to-introduce-physical-education-for-schoolgirls-idUSKBN19W1DR/>.

"Top Indian female wrestler quits in protest over new president of wrestling body." *Reuters*.
<https://www.reuters.com/world/india/top-indian-female-wrestler-quits-protest-over-new-president-wrestling-body-2023-12-21/>.

Roper, Gemma. "Juffali: 'Motorsport and racing were an impossible dream.'" *Females in Motorsport*
<https://www.femalesinmotorsport.com/post/reema-juffali-motorsport-and-racing-were-an-impossible-dream>

Saina. Directed by Amole Gupte. 2021.

Sambidge, Andy. "Saudi Arabia's next goal: women's sport." *Arabian Gulf Business Insight*.
<https://www.agbi.com/business-of-sport/2022/08/saudi-arabias-next-goal-womens-sport/>

Sharma, Sanjay. "Saina Nehwal: Indian badminton's success story." *Al Jazeera* (Doha).
<https://www.aljazeera.com/sports/2015/6/1/saina-nehwal-indian-badminton-success-story>.

Specia, Megan. "Saudi Arabia Granted Women the Right to Drive. A Year on, It's Still Complicated." *The New York Times*.
<https://www.nytimes.com/2019/06/24/world/middleeast/saudi-driving-ban-anniversary.html>.

Vasudevan, Shyam. "2023: the year wrestling changed everything, and yet things remained the same." *ESPN*.
https://www.espn.com/wrestling/story/_/id/39186019/wfi-protest-vinesh-phogat-sakshi-malik-bajrang-punia-brij-bhushan-sharan-singh-story-2023.

"Vision 2030." Vision 2030. <https://www.vision2030.gov.sa/en>.

"What's the secret weapon keeping girls in school in India?" *WE Charity*.
<https://www.we.org/en-US/we-stories/global-development/clean-water-sends-girls-to-school-in-india>.

Zambarakji, Angie. "Saudi Ban on Women's Sports Blamed for Rising Obesity." *The Bureau of Investigative Journalism*.
<https://www.thebureauinvestigates.com/stories/2012-05-18/saudi-ban-on-womens-sports-blamed-for-rising-obesity/>.

Constraining Observables in Neutron Stars with Varying Equations of State

By Sahil Arora

1. Abstract

In this paper, we encompass the existing work on equations of state of neutron stars from existing literature and establish the groundwork with mass-radius curves which are used to find the corresponding maximum mass with minimum radius across the Equations of State (EOSs). We use SLy, WFF1, H7, BBB2, and the APR 4 EOS and primarily focus on the properties of SLy. Using the results from MR curves, we calculate and, in the process, set maximum bounds on the spin of a NS and also on the energy loss rate as calculated assuming dipolar radiation from a NS with misaligned magnetic and rotation axes purely due to its rotational kinetic energy. We then proceed to discuss the reasons for the observed properties, that is nucleon-nucleon interaction and introduce more calculations and compare with observed data from various databases.

Keywords

EOS, Neutron Stars, Spin period, Energy Emission, MR Curve

2. Introduction

Neutron Stars (NS) are compact objects that are formed after a collapse of a massive star in a supernova explosion. NS are formed from a core collapse supernova (CCSN), leaving behind a dense and compact core of the Proto-Neutron Star (PNS). They are the densest stars in the universe and get their name because they are made of neutrons. These stars provide laboratories to explore high-density nuclear matter by using a well-suited equation of state (EoS) relating the density and pressure of matter. Consequently, it gives the relation between the mass and radius of the star and describes tidal deformability- the deformation of a celestial body caused by the gravitational forces of other celestial bodies- and moment of inertia- the ability of a body to resist angular acceleration [1-2].

Neutron stars are the most compact known objects without event horizons and therefore serve as extraordinary laboratories for dense matter physics. The internal composition of the cores of neutron stars is currently poorly understood. [3]

The study of internal structures of NS is possible by formulating a theoretical model, which are theoretical EOSs and the realistic nature of EoS can be understood with the help of astrophysical measurements. These provide macroscopic properties of neutron stars and give information about the softness, which is when pressure rises slowly and consistently with increase in density and stiff-ness, which is when pressure increases rapidly with slight changes in density, of EoS, pressure and density. Einstein's theory of relativity plays a significant role in describing the physical properties of neutron stars. Anglo-Indian astrophysicist Chandrasekhar considered the degenerate matter and fixed a theoretical upper limit of 1.44 solar masses. It conserves the stability of non-rotating neutron stars [4]. This limit of maximum mass does not match with the results of

massive pulsars which are rapidly rotating NS with more emissions in x-rays and radio waves [5-9]. Several attempts have been put to obtain constraints on the mass and radius relation [10-13].

For soft EoS, the value of maximum mass and radius are up to 1.4 M and 10 km respectively, while for very stiff EoS, these could have values up to 1.8 and 15 km, respectively. [2]. Further astrophysical scenarios characterized by high nuclear density and large temperature, and where the EoS plays a relevant role are (i) core-collapse supernova explosions (CCSN), which lead to the formation of a very hot proto-neutron star (PNS) and subsequently to a cold NS or to a black hole (BH) [14-17], (ii) merging of NS in close binary systems, NS–NS and NS–BH [18-20]. The dynamical evolution and the structure of the forming final compact object are strongly determined by the nuclear-matter EoS, as well as the nucleosynthesis conditions and the emitted neutrino spectra. Therefore there is a deep connection between the gross features of these astrophysical phenomena and the underlying microphysics ingredients which arise from the interactions among the particles. [21]

In the Methods section of this paper we introduce the equations we shall use throughout the study and provide the required derivations. Within the Methods section we explore mass-radius curves and move on to constraining the spin of NS and the energy loss. We then discuss the results and why they happen, exploring the workings of the EOSs with its constituent particles and conclude with a comparative study with real life observations.

3. Methods

3.1 Equations Of State

The EoS depends primarily on the matter constituents and how these constituents interact with each other, which leads us to the observed density of the neutron stars. We use $\rho_0 = 2.4 \times 10^{14} \text{ g cm}^{-3}$. This paper discusses SLy, WFF1, H7, BBB2, and APR 4 EOSs.

3.1.1 Skyrme Lyon

The SLy (Skyrme Lyon) models—due to the emphasis put on their neutron-excess dependence—seem to be particularly suitable for the calculations of the properties of neutron-star crust. A new set of the Skyrme-type effective N–N interaction, based on an approach which may be more appropriate, as far as the applications to a very neutron-rich matter are concerned.

Relevant additional experimental items concerning neutron rich nuclei (isovector effective masses), constraints of spin stability, and requirement of consistency with the UV14+VIII equation of state (EOS) of dense neutron matter of Wiringa et al. [22] for $\rho_0 < \rho < 1.5 \text{ fm}^{-3}$ [23] Further in the paper, we discuss the properties obtained for the SLy equation along with other parameters.

3.1.2 WFF1

The Wiringa–Fiks–Fabrocini (WFF1) EOS calculations are based on a Hamiltonian that fits N–N scattering data, the properties of deuterons, and includes explicit forms of three-nucleon

interaction to obtain a fit to nuclear matter equilibrium properties. The WFF1 is with interaction terms AV14+U VII [22, 24]

3.1.3 H7

The H7 EOS consists of nucleons and hyperons (discussed below), making it different from the other EOSs discussed. For small strangeness fractions ($f \leq 1$), matter is dominated by NLJ composition and the calculated binding energy closely follows that calculated by using the hyperon potentials of our previous calculations. For larger strangeness fractions ($f \geq 1$), the calculated binding energy increases substantially beyond any previous calculation due to a phase transition into NSJ dominated matter. To calculate the model, a relativistic mean-field (RMF) model is used [25]. These properties and results are further discussed below.

3.1.4 BBB2

The EOS with which we have calculated employs the Brueckner–Hartree–Fock model EOS: BBB2. This EOS is derived from the Brueckner-Bethe-Goldstone many-body theory. The EOS, primarily employed in three-body interaction (discussed below), predicts a higher maximum mass for neutron stars [26].

3.1.5 APR 4

Properties of dense nucleon matter and the structure of neutron stars are studied using variational chain summation methods and the new Argonne AV14 two-nucleon interaction, which provides an excellent fit to all of the nucleon-nucleon scattering data. The neutron star's gravitational mass limit obtained with this interaction is 1.67 M. Boost corrections to the two-nucleon interaction, which gives the leading relativistic effect, as well as three-nucleon interactions, are also included in the nuclear Hamiltonian. Their successive addition increases the mass limit to 1.80 and 2.20 M_s [27]. More about its N-N interaction is discussed below.

3.2 Mass-Radius Curve

The Tolman-Oppenheimer-Volkoff equation (hereby referred to as TOV) is an important component to get the desired mass-radius (M-R) curves. The TOV [28,29] equation can be derived by positing a generic metric valid for a static isotropic spacetime and assuming a perfect fluid for the stress–energy tensor. If we assume spherical symmetry, implying that the pressure P and density ρ are functions of only the radial coordinate r then one obtains the TOV equation,

$$dP/dr = - G/r^2 (M(r) + 4\pi r^3 P(r)/c^2) (\rho(r) + P(r)/c^2) (1 - 2GM(r)/c^2 r)^{-1}$$

This equation represents the reduction of Einstein's equations for the interior of a spherical, static object. This equation is the GR analog of the Newtonian hydrostatic condition, given as,

$$dP/dr = - GM\rho/r^2$$

The TOV equation is routinely used to determine the mass–radius relation of neutron stars as well as the maximum mass, once an equation of state $P(\rho)$ is specified. The main difference between the TOV equation and its Newtonian counterpart is that in the latter, pressure supports the object against gravity, whereas in the TOV equation, pressure contributes to the gravitational field. That is why, in very strong gravitational fields, pressure expedites gravitational collapse, ultimately causing the object to collapse to a black hole [30].

The Mass-Radius Curve (hereafter referred to as MR), is the first step in the evaluation. Integration of the TOV equation (2.1) gives us the desired MR curve which highlights the differences that exist between different EOS. In contrast to Newtonian gravity, a maximum mass for each EOS exists in GR. In addition, a minimum mass exists, but is only about $0.09 M_s$ [31]. Curve of the graph helps us (i) Find the regions of stiffness and softness of the EoS (ii) Find the radius for a particular mass allowed by the EoS which will be used further in the paper. The integration, done using Python [32]. It is from the center to the surface where $\rho_s = 0$ and $\rho_c = \infty$, which sets the boundary conditions and the parameters are set by the EoS. A majority of celestial objects are thought to exist in a vacuum or nearly vacuum spacetime. At the interface between these two regions, there is a suitable connection between the inner geometry (M) and the exterior vacuum area (M+). The optimal solution for the outside in this scenario is provided by an accurate Schwarzschild spacetime. It is a solution to Einstein's GT for a non-rotating, non-charged mass [33]

$$ds^2 = ((2M/r) - 1)dt^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) + dr^2/(1 - 2M/r)$$

where M is the mass of the central object. The first term describes the time component, the second covers the angular component and the last term covers the radial component. It is used in describing the Schwarzschild radius.

3.3 Spin Period

The maximum spin period of a neutron star can be given by the Mass-Shedding Frequency f_s . This is the frequency at which the gravitational attraction of the star is overpowered by the centrifugal force.

To calculate the same, we set the gravitational acceleration equal to the centripetal acceleration, given by $GM/R^2 = R\Omega^2$ where G is the Newtonian gravitational constant, M is the mass of the star, R is the corresponding radius (seen in relation with the MR curve), and Ω is the angular frequency. The equation can further be simplified as

$$\sqrt{GM/R^3} = \Omega$$

The calculation done is with a version of the actual workings that is based on classical mechanics and does not include GR. This model accounts for the maximum angular frequency possible with a given mass and its radius, not accounting for the dynamics of individual EOS. However, we shall also use the rotating neutron star code (rns) [34] for advanced calculation.

S.No.	EOS	$M_{\text{max}}(M_s)$	$R_{(M_{\text{max}})}(\text{Km})$	$R_{1.4}(\text{Km})$	$f_s(\text{KHz})$	$f_m(\text{KHz})$
1	SLy	2.049	10.06	11.736	2.60	1.81
2	WFF1	2.133	9.6	10.414	2.85	2.04
3	H7	1.638	11.47	12.91	1.94	1.41
4	BBB2	1.900	9.5	11.139	2.74	1.90
5	APR 4	2.213	10.37	11.43	2.58	1.94

Table 1: Table with values evaluated using data from [36-39]; Shows the difference in frequencies across the EOS for the relative stiffness

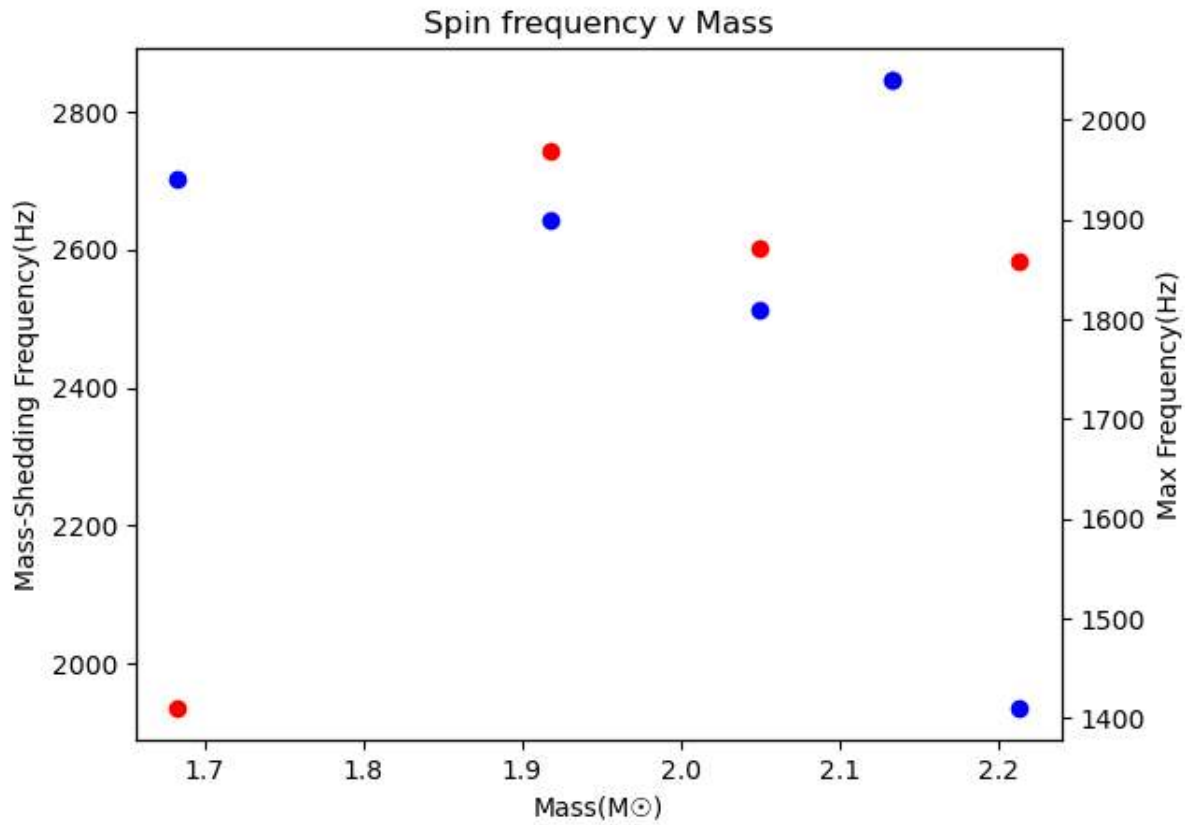


Figure 1: The graph shows a rough distribution of the values of mass-shedding frequency against mass and maximum frequency values obtained from [35]

3.4 Radiative Energy Loss

It should come out as no surprise that neutron stars do not retain all the energy and radiate. In this paper we shall assume in the case of a pulsar that all radiation takes place at the magnetic dipoles due to its rotational kinetic energy using the magnetic dipole model [40].

In the oblique rotator version of the model, we shall assume that a neutron star in vacuum rotates with angular frequency Ω_s (sec. 3.3) and has a magnetic dipole m at an angle α to the rotation axis [40].

From [40], it is seen that $|m| = B_p R^3 / 2$ where B_p is the magnetic field at the dipoles and R is the radius of the star. In the calculations, we will be using a standard value for $B_p = 10^{12} \text{G}$

Further in calculations, we find

$$\dot{E} = - B_p^2 R^6 \Omega^4 \sin^2 \alpha / 6c^3$$

Where E is the energy lost over time. If we use I , the moment of inertia of the star, the above eq can be written as $\dot{E} = I \Omega \dot{\Omega}$

since the radiated energy comes from the kinetic energy of the star, given as $E = (1/2) I \Omega^2$ which is the spin-down luminosity In calculations we shall use $I \approx 10^{45} \text{g cm}^2$ (see table), $B_p \approx 10^{12} \text{G}$ for a pulsar, $\approx 10^8 \text{G}$ for a millisecond pulsar, and $\approx 10^{14} \text{G}$ for a magnetar which is a neutron star with very strong magnetic fields, and $\sin^2 \alpha = 1$ (for maximum energy loss) and $\sin^2 \alpha = 0.1170$ ($\alpha = 20^\circ$) [41].

The calculations using the above equations for the energy lost are discussed in the Results section.

4 Results and Discussion

In this section we shall discuss and group the results obtained from the previous sections and also elaborate on how good of a fit each of the models are (see Figure 2).

S.No.	EOS	$M_{\text{max}}(M_s)$	$R_{(M_{\text{max}})}(\text{Km})$	$I \text{ (g cm}^2\text{)}$
1	SLy	2.049	10.06	1.649×10^{45}
2	WFF1	2.133	9.6	1.564×10^{45}
3	H7	1.638	11.47	1.761×10^{45}
4	BBB2	1.900	9.5	1.377×10^{45}
5	APR 4	2.213	10.37	1.893×10^{45}

Table 2: The above table displays the maximum moment of inertia possible with different EOS, measured in cgs units.

4.1 MR Curve and N-N interaction

Continuing section 3, the SLy and APR 4 predict a stiff neutron star, BBB2 and WFF1 predict a relatively softer neutron star, whereas, H7 predicts a soft star. This can very well be seen in

the MR curve where stiffer equations predict a higher possible mass and a smaller possible radius. Secondly, stiffer equations have a steeper slope.

The above parameters classify the stiffness of an EOS. It is to note the reasons because of which they arise: the nucleon-nucleon(N-N) interaction. SLy and APR 4 N-N interaction incorporates repulsion at high densities $\rho \geq 10^{14} \text{ g cm}^{-3}$.

Conversely, the WFF1, H7 and BBB2 N-N interactions incorporate attraction at higher densities. In likeness there are differences, however. The models each use a different basis for their N-N interactions as given in Table 3.

S.No.	EOS	Type	N-N Interaction Type	N-N Constituents
1	SLy	Stiff	Repulsive	p,n
2	WFF1	Soft	Attractive	p,n
3	H7	Soft	Attractive	p,n, Λ , Ξ , Σ
4	BBB2	Soft	Attractive	p,n
5	APR 4	Stiff	Repulsive	p,n

Table 3: Summarizes the type of EOS and its N-N interaction types along with its constituents

The difference in how the nucleons are parameterized in each EoS gives us the difference in the results we see. The SLy N-N interaction belongs to the Skyrme effective interactions, which are forces that are primarily used within Hartree-Fock or Density Functional Theory (DFT). This model of interaction is effective in describing rich matter effectively. Features of the SLy N-N interactions include its terms that depend on density, including spin-orbit coupling and the energy density. The repulsion seen in nucleons in SLy is by Pauli exclusion principle and provides the base for further modeling.

It is seen that SLy primarily models the n and p interactions (Two-Body) and limits itself in three-body interactions, thus, not including hyperons (see table 4). Under the conditions prevailing in the neutron star interior, both nucleon phases are immersed in an essentially homogeneous electron gas. The uniformity of the electron gas implies that – in the bulk approximation – its presence does not influence the equilibrium between the “i” and “o” phases of the nucleon component of the matter, ensured by the (strong) N-N interaction. [42] [43]

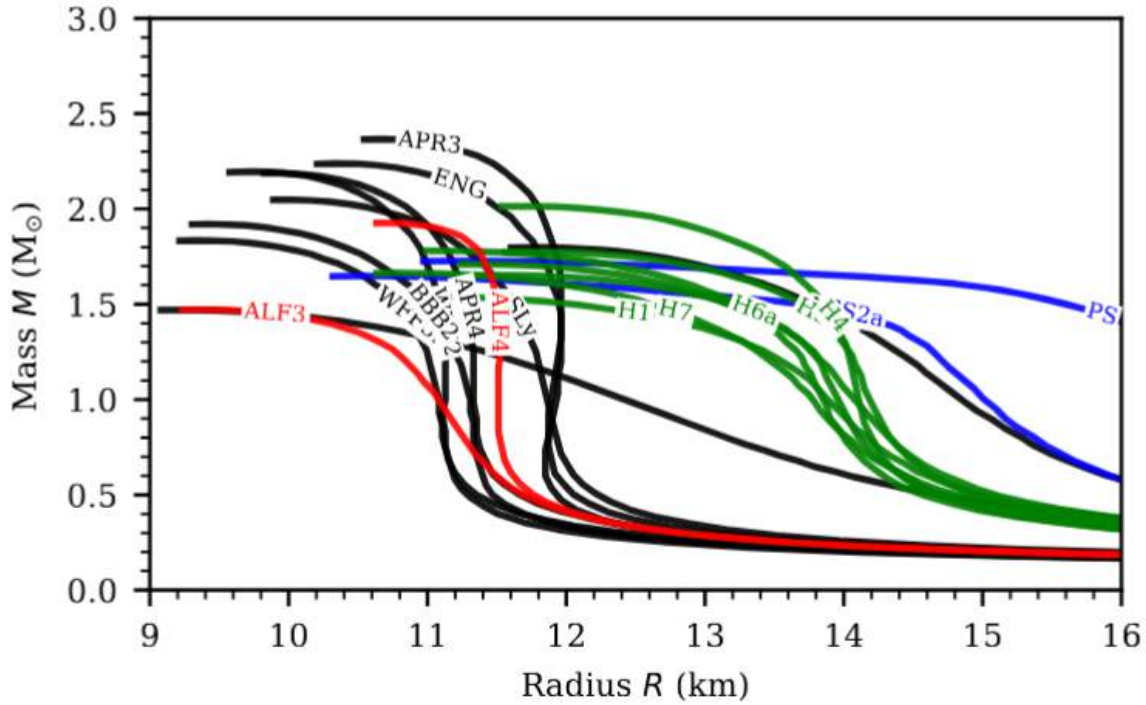


Figure 2: The graph employs 20 different EoS varying in constituents and parameters. Black lines refer to EoS with particles as npem, Red referring to quark, Green indicating hyperons and Blue being mesons.

For WFF1, The EOS used does not include effects of strange baryons at supra-nuclear densities. The influence of hyperons on the nuclear EOS has been recently investigated [24] The APR 4 EOS uses a three-nucleon potential in addition to an Argonne 18 potential, and its variant the WFF1 EOS, which differs on the nucleon potential model used. [44] The difference between the energies predicted by A18+dv and A18+dv+UIX* Interactions at higher densities are due to three-body forces; it is smaller than that between the energies obtained from A18 and A18+UIX by almost a factor of two. [45].

At low density $0.2 < n < 0.4 \text{ fm}^{-3}$, the EOS of Wiringa, Fiks, and Fabrocine (WFF) is the softest, while for high densities it is the stiffest. Wiringa, Fiks, and Fabrocine attribute the softening at low densities to neutral pion condensation. This additional softening allows for rapid enough rotation, while the stiffening at high densities (in this case to the causal limit) provides for a maximum mass significantly above $1.44 M_s$ [46].

The H7 model is an effective model where the parameters are adjusted to the known properties of nuclei and hyper-nuclei. It includes in our extended RMF model all the $1/2^+$ baryons of the lowest SU(3) flavor octet, as well as hidden-strangeness meson exchange to allow for possibly strong hyperon-hyperon (YY) interactions. While it is true that the RMF model 2 is a schematic model

Terms	Parameter Modeled
Central Potential	Attractive and Repulsive Forces in N-N interaction
Density-Dependence	Accounts for the exclusion principle, adjusts forces on regional density
Kinetic Energy	Models for the motion of nucleons focusing on their wave functions and momenta
Spin-Orbit Coupling	Capture the interaction between N spin and orbital motion
Symmetry Energy	Represents energy cost of creating imbalance between n and p in nuclear matter SLy has a strong dependence on this energy in extreme n:p ratios, impacting the stiffness.
Exchange	Correct the interaction potential that arises due to asymmetry in the model and help describing the quantum nature
Three-Body Interaction	Effective term in Skyrme potential to mimic three-body interactions
Coulomb	Accounts for the electromagnetic repulsion between positively charged p, which are a negligible fraction in a NS

Table 4: The table discusses the terms that are included in the Skyrme Lyon EoS and the parameters they account for, creating the environment that has been discussed.

and is linked only indirectly to the underlying baryon-baryon interactions, it is nevertheless constrained by Λ and Ξ nuclear phenomenology, and by the few $\Lambda\Lambda$ hyper-nuclear species [25].

BBB2 uses the Argonne AV14 and the Paris two-body nuclear force, implemented by the Urbana model for the three-body force, which gives us the observed M_{max} . In BBB2, two different two-body potentials have been used, the Argonne AV14 and the Paris interactions. The latter turns out to produce a slightly stiffer EOS already at two-body level. The same three-body force, within the Urbana model, has been added to both two-body interactions, and this produces similar saturation points close to the empirical one [47].

4.2. Spin frequencies

Other than their mass-shedding frequencies, NS also has a Keplerian angular frequency. Here, we shall use one for a non-rotating NS with a maximum mass

$$\Omega_k = 7.7 \cdot 10^3 (M_{\text{max}}/M_s)^{1/2} (R_{\text{max}}/10 \text{ Km})^{-3/2} \text{ s}^{-1}$$

Here the subscripts “max” denote the maximum mass non-rotating star of a given equation of state, and M_{max} is the gravitational mass. Some motivation for this simple relationship comes

from the so-called Roche model [40] in Newtonian gravity. This model assumes that the distribution of the bulk of the mass is unchanged by rotation, which, for highly centrally condensed stars, the ones with the largest Keplerian rates, is a good approximation. In this model the Keplerian rate for a star of mass M and non-rotating radius R is [46].

$$\Omega_k = 6.3 \cdot 10^3 (M_{\text{max}}/M_s)^{1/2} (R_{\text{max}}/10 \text{ Km})^{-3/2} \text{ s}^{-1}$$

EOS	Frequency (Hz)	$\dot{E}$ (10^{43} ergs s^{-1}) [$\alpha=90^\circ$]	$\dot{E}$ (10^{43} ergs s^{-1}) [$\alpha=20^\circ$]
Sly	1810	10.727	1.255
WFF1	2040	13.071	1.529
H7	1410	8.676	1.015
BBB2	1900	9.235	1.081
APR 4	1940	16.979	1.987

Table 5: Maximum energy dissipation at magnetic dipoles of a neutron star due to its rotation, giving a direct relation between angular frequency and energy emitted

The results of mass-shedding frequency in table 1 uses a non-relativistic and does not consider the properties of the EOS and thus, the maximum frequency shall be used. The calculations show a vast difference in Mass-shedding and maximum frequency suggesting any direct approach wouldn't be appropriate and the relativistic effects have to be considered. As high the values are in either case, it has not been observed.

The discovery of PSR J1748-2446ad, a 716-Hz (1.396 ms) eclipsing binary radio pulsar in the globular cluster Terzan 5, using the Green Bank Telescope (GBT). It is the fastest-spinning pulsar ever found, breaking the 23-year-old record held by the 642-Hz pulsar B1937+21.

It however, does not limit the possibilities of pulsars with higher frequency, because PSR J1748-2446ad is eclipsed for 40% of its orbit and has a low flux density ≈ 0.08 mJy at 2 GHz .

These properties make detection difficult and suggest that even faster-spinning neutron stars exist, but their strong winds are ablating their companions, causing long eclipses of the pulsar. If the pulsar has a mass less than $2 M_s$, then its radius is constrained by the spin rate (independent of the true equation of state) to be ≤ 16 Km. [48]

4.3 Energy Emissions from dipoles

For a pulsar with a given angular frequency with different modeled EOS, the energy emitted can be estimated to be (table 5).

The values of Ω and E can be used as the maximum possible values as they are calculated at the maximum possible state and thus, $\Omega = \Omega_i$ and $E = E_i$, where i stands for initial.

When a time parameter is introduced into the Energy emission formula, it is seen to have a peak when the star is newly formed from a PNS and the emission decreases as the pulsar ages, as is seen in the frequency curve. To evaluate the same, we use the steps followed in [40]

$$T = 6Ic^3/B_p^2 R^6 \sin^2 \alpha \Omega_0^2$$

where T is a defined characteristic age. This can further be utilized in

$$\Omega = \Omega_i (1 + (2\Omega_i t)/\Omega_0 T)^{-1/2}$$

Where t is the current age left as a free parameter, we obtain the plot of angular frequency over time as above

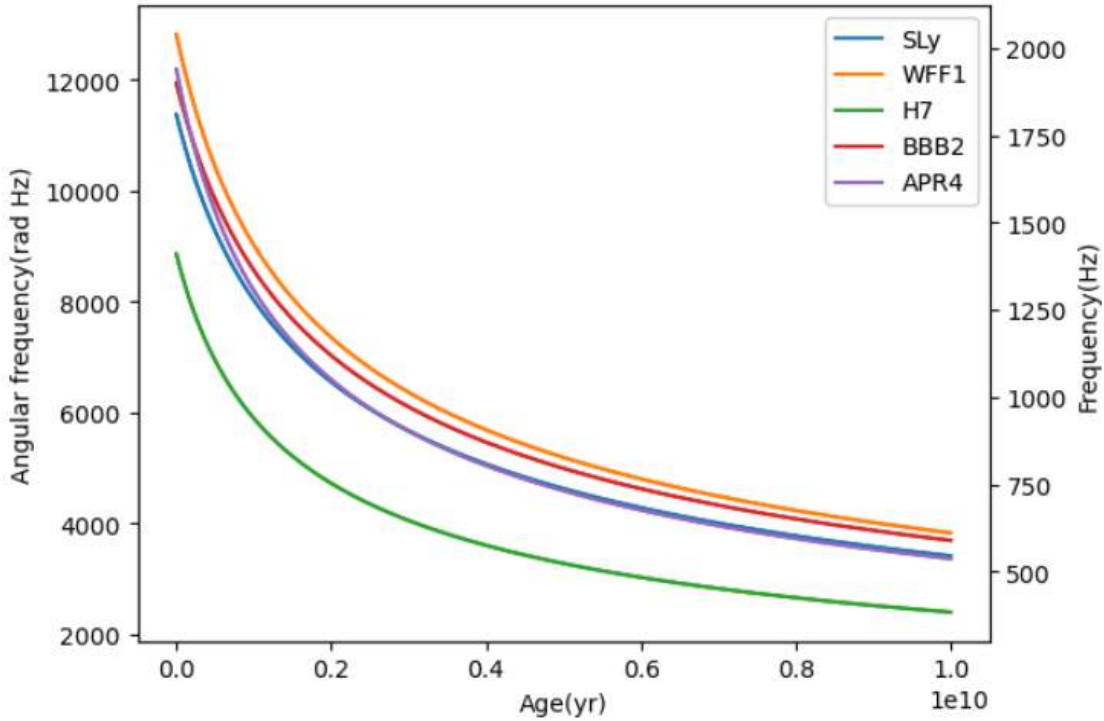


Figure 3: Frequency-Time Curve

4.4 Observed Data

The most commonly seen NS do not spin very rapidly as is seen from table 6 where most NS are seen to be second pulsars and traces of millisecond pulsars. The observed masses of NS also vary vastly. Most NS are seen in the ranges of 1.2 and 1.6 M and very few reach 2 M and close to 1 M. This makes a NS of 1.4 M more likely, as seen.

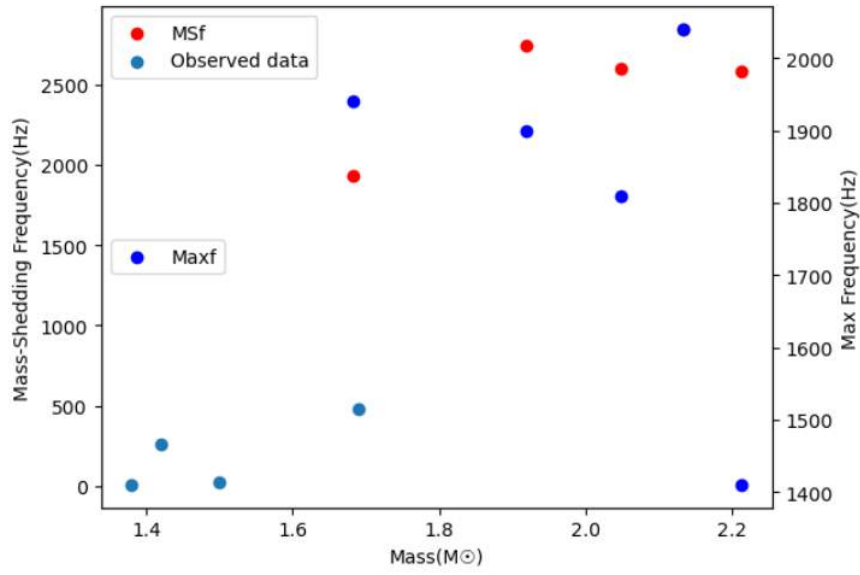


Figure 4: Real V Theoretical Frequencies

No.	Pulsar Name	Period (s)	Frequency (Hz)
1	J0006+1834	0.693	1.443
2	J0011+08	2.552	0.392
3	J0023+0923	0.0030	333.334
4	J0050+03	1.366	0.732
5	J0058+4950	0.996	1.004
6	J0103+54	0.354	2.825
7	J0107+1322	1.197	0.835
8	J0122+1416	1.389	0.720
9	J0139+3336	1.248	0.801
10	J0146+31	0.938	1.066
11	J0229+20	0.806	1.241
12	J0241+16	1.544	0.648
13	J0244+14	2.127	0.470
14	J0302+2252	1.207	0.829
15	J0329+1654	0.893	1.120

Table 6: This table [49-52] has the observed time periods of certain pulsars and their angular frequency has been calculated using $f=1/T$

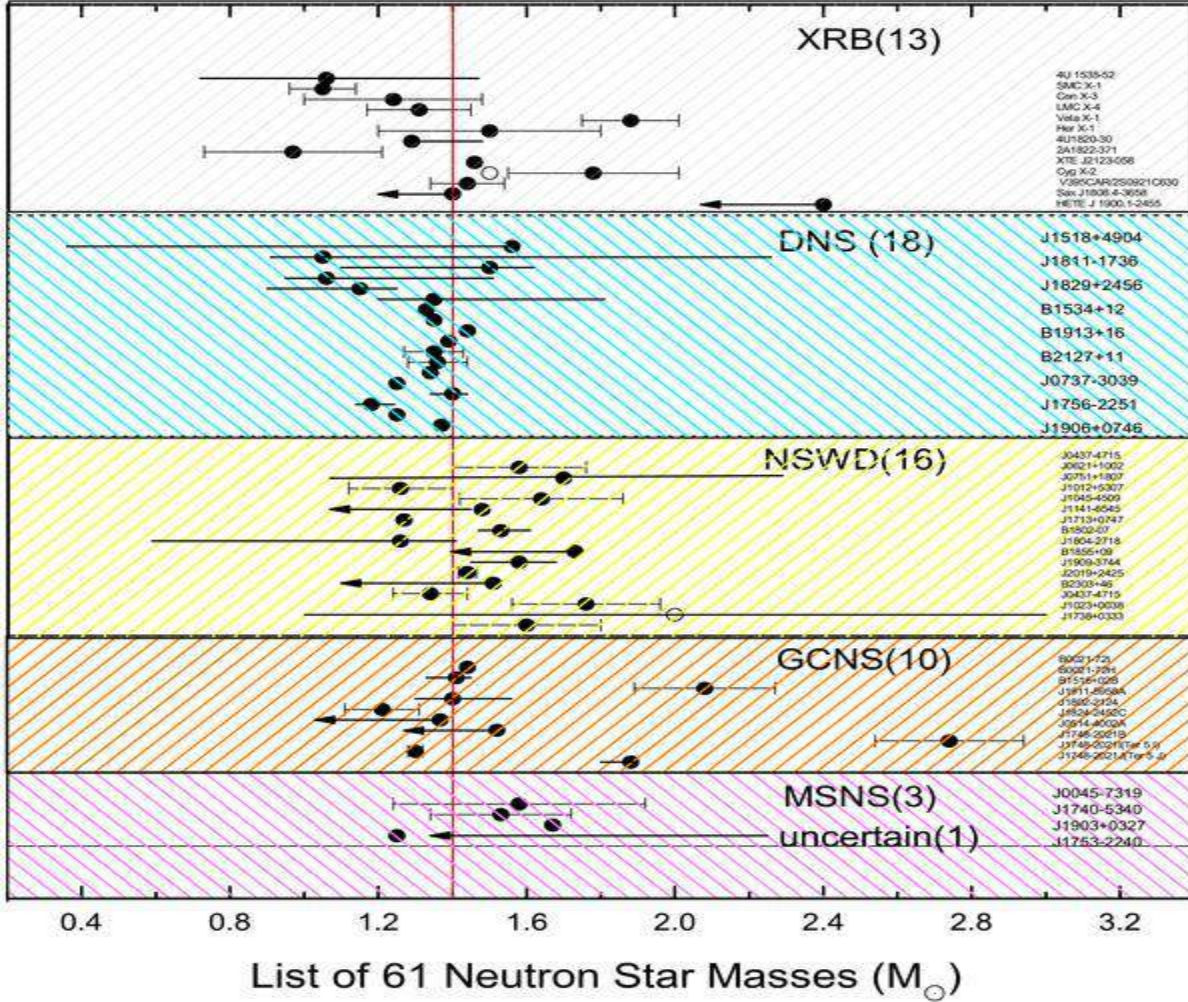


Figure 5: Observed Neutron Star Masses taken from Zhang et. al [53]

5. Conclusion

In this paper we went through the most widely used EOS in the field and found the observables that depend on the EOS. In our analysis, we found out the relation between the MR Curve and the N-N interaction that alters it. We also set maximum limits on the angular frequency of NS that is seen to vary according to the EOS. It was also observed that for stiffer EOS, a higher maximum angular frequency is seen.

Using the calculations of the maximum spin period, we also set upper bounds on the maximum energy loss of a NS through its magnetic dipoles solely based on its rotational kinetic energy with two variations in angle of incline. The energy lost and the angular frequency also helped us set the maximum moment of inertia of an EOS. Setting a time parameter to the equation that

accounts for the age of the pulsar gives us a plot that explains that as a NS ages and slows down, its energy loss decreases exponentially over a large period of time.

This paper dived deeper into Skyrme Lyon (SLy) as we believe it matches closely with observed data and proposes further enhancement on the model from SLy 7. We also urge the need to focus more on the theories for EOS of NS and also more missions like NICER to be suggested

Works Cited

- [1] A. W. Steiner, S. Gandolfi, F. J. Fattoyev, and W. G. Newton. Using neu-tron star observations to determine crust thicknesses, moments of inertia, and tidal deformabilities. *Phys. Rev. C*, 91:015804, Jan 2015.
- [2] Nisha Godani and Gauranga C. Samanta. Neutron stars with polytropic equation of state. *New Astronomy*, 107:102148, 2024.
- [3] James M. Lattimer and Madappa Prakash. Neutron star observations: Prognosis for equation of state constraints. *Physics Reports*, 442(1):109–165, 2007. The Hans Bethe Centennial Volume 1906-2006.
- [4] S. Chandrasekhar. The Maximum Mass of Ideal White Dwarfs. , 74:81, July 1931.
- [5] Feryal Ozel and Paulo Freire. Masses, radii, and the equation of state of neutron stars. *Annual Review of Astronomy and Astrophysics*, 54:401–440, 2016.
- [6] P. Demorest, T. Pennucci, S. Ransom, M. S. E. Roberts, and J. W. T. Hessels. A two-solar-mass neutron star measured using shapiro delay. *Nature*, 467:1081–1083, 2010.14
- [7] Paulo C. C. Freire, A. Wolszczan, M. van den Berg, and J. W. T. Hessels. On the nature and evolution of the unique binary pulsar j1903+0327. *The Astrophysical Journal*, 675:670–680, 2008.
- [8] P. Haensel and A. Y. Potekhin. Analytical representations of unified equations of state of neutron-star matter. *Astronomy Astrophysics*, 380:151–167, 2001.
- [9] A W Steiner, C O Heinke, S Bogdanov, C K Li, W C G Ho, A Bahramian, and S Han. Constraining the mass and radius of neutron stars in globular clusters. *Monthly Notices of the Royal Astronomical Society*, 476(1):421–435, 01 2018.
- [10] I. Wasserman and S. L. Shapiro. Masses, radii, and magnetic fields of pulsating X-ray sources : is the “standard” model self-consistent ? , 265:1036–1046, February 1983.
- [11] John Antoniadis, Paulo C. C. Freire, Norbert Wex, Thomas M. Tauris, Ryan S. Lynch, Marten H. van Kerkwijk, Michael Kramer, Cees Bassa, Vik S. Dhillon, Thomas Driebe, Jason W. T. Hessels, Victoria M. Kaspi, Vladislav I. Kondratiev, Norbert Langer, Thomas R. Marsh, Maura A. McLaughlin, Timothy T. Pennucci, Scott M. Ransom, Ingrid H. Stairs, Joeri van Leeuwen, Joris P. W. Verbiest, and David G. Whelan. A massive pulsar in a compact relativistic binary. *Science*, 340(6131):1233232, 2013.
- [12] Roger W. Romani, Alexei V. Filippenko, Jeffrey M. Silverman, and S. Bradley Cenko. Psr j13113430: A heavyweight neutron star with a flyweight helium companion. *The Astrophysical Journal Letters*, 760:L36, 2012.
- [13] M. H. van Kerkwijk, R. P. Breton, and S. R. Kulkarni. Evidence for a massive neutron star from a radial-velocity study of the companion to the black-widow pulsar psr b1957+20. *The Astrophysical Journal*, 728:95, 2011.
- [14] Anthony Mezzacappa. ASCERTAINING THE CORE COLLAPSE SUPERNOVA MECHANISM: The state of the art and the road ahead. *Annu. Rev. Nucl. Part. Sci.*, 55(1):467–515, December 2005.
- [15] Hans-Thomas Janka. Explosion mechanisms of core-collapse supernovae. *Annu. Rev. Nucl. Part. Sci.*, 62(1):407–451, November 2012.

- [16] Madappa Prakash, Ignazio Bombaci, Manju Prakash, Paul J. Ellis, James M. Lattimer, and Roland Knorren. Composition and structure of protoneutron stars. *Physics Reports*, 280(1):1–77, 1997.
- [17] J A Pons, S Reddy, M Prakash, J M Lattimer, and J A Miralles. Evolution of proto–neutron stars. *Astrophys. J.*, 513(2):780–804, March 1999.
- [18] Masaru Shibata and Keisuke Taniguchi. Coalescence of black hole–neutron star binaries. *Living Rev. Relativ.*, 14(1):6, August 2011.15
- [19] Stephan Rosswog. The multi-messenger picture of compact binary mergers. *International Journal of Modern Physics D*, 24(05):1530012, 2015.
- [20] Luca Baiotti and Luciano Rezzolla. Binary neutron star mergers: a review of einstein’s richest laboratory. *Rep. Prog. Phys.*, 80(9):096901, September 2017.
- [21] G.F. Burgio, H.-J. Schulze, I. Vidaña, and J.-B. Wei. Neutron stars and the nuclear equation of state. *Progress in Particle and Nuclear Physics*, 120:103879, 2021.
- [22] R. B. Wiringa, V. Fiks, and A. Fabrocini. Equation of state for dense nucleon matter. *Phys. Rev. C*, 38:1010–1037, Aug 1988.
- [23] F Douchin and P Haensel. Inner edge of neutron-star crust with sly effective nucleon-nucleon interactions. *Physics Letters B*, 485(1):107–114, 2000.
- [24] Manju Prakash, E. Baron, and Madappa Prakash. Rotation of stars containing strange quark matter. *Physics Letters B*, 243(3):175–180, 1990.
- [25] Jürgen Schaffner-Bielich and Avraham Gal. Properties of strange hadronic matter in bulk and in finite systems. *Physical Review C*, 62(3), August 2000.
- [26] M. Baldo, I. Bombaci, and G. F. Burgio. Microscopic nuclear equation of state with three-body forces and neutron star structure, 1997.
- [27] A. Akmal, V. R. Pandharipande, and D. G. Ravenhall. Equation of state of nucleon matter and neutron star structure. *Phys. Rev. C*, 58:1804–1828, Sep 1998.
- [28] Richard C. Tolman. Static solutions of einstein’s field equations for spheres of fluid. *Phys. Rev.*, 55:364–373, Feb 1939.
- [29] J. R. Oppenheimer and G. M. Volkoff. On massive neutron cores. *Phys.Rev.*, 55:374–381, Feb 1939.
- [30] Sajal Gupta and Shantanu Desai. Galaxy cluster hydrostatic masses using tolman–oppenheimer–volkoff equation. *Physics of the Dark Universe*, 28:100499, 2020.
- [31] James M. Lattimer. Neutron star equation of state. *New Astronomy Reviews*, 54(3):101–109, 2010. *Proceedings: A Life With Stars*.
- [32] GitHub - natj/tov: TOV solver for neutron stars — github.com. <https://github.com/natj/tov>.
- [33] S.K. Maurya, Abdelghani Errehymy, Ksh. Newton Singh, Orhan Donmez, Kottakkaran Sooppy Nisar, and Mona Mahmoud. Self-bound isotropic models in $f(q)$ gravity and effect of $f(q)$ parameter on mass–radius relation and stability of compact objects. *Physics of the Dark Universe*, 46:101619, 2024.16
- [34] Manual.Pdf at main · cgca/rns.

- [35] Jocelyn S Read, Benjamin D Lackey, Benjamin J Owen, and John L Friedman. Constraints on a phenomenologically parameterized neutron-star equation of state. December 2008.
- [36] Jocelyn S. Read, Benjamin D. Lackey, Benjamin J. Owen, and John L. Friedman. Constraints on a phenomenologically parametrized neutron-star equation of state. *Physical Review D*, 79(12), June 2009.
- [37] S. Typel, M. Oertel, and T. Klähn. Compose compstar online supernova equations of state harmonising the concert of nuclear physics and astrophysics compose.obspm.fr. *Physics of Particles and Nuclei*, 46(4):633–664, Jul 2015.
- [38] M. Oertel, M. Hempel, T. Klähn, and S. Typel. Equations of state for supernovae and compact stars. *Rev. Mod. Phys.*, 89:015007, Mar 2017.
- [39] S. Typel, M. Oertel, T. Klähn, D. Chatterjee, V. Dexheimer, C. Ishizuka, M. Mancini, J. Novak, H. Pais, C. Providencia, A. Raduta, M. Servillat, and L. Tolos. Compose reference manual, 2022.
- [40] Stuart L. Shapiro and Saul A. Teukolsky. Black holes, white dwarfs and neutron stars. The physics of compact objects. 1983.
- [41] Scott Ransom. Basic radio pulsar astronomy, 2017.
- [42] F Douchin, P Haensel, and J Meyer. Nuclear surface and curvature properties for sly skyrme forces and nuclei in the inner neutron-star crust. *Nuclear Physics A*, 665(3):419–446, 2000.
- [43] E. Chabanat, P. Bonche, P. Haensel, J. Meyer, and R. Schaeffer. A skyrme parametrization from subnuclear to neutron star densities part ii. nuclei far from stabilities. *Nuclear Physics A*, 635(1):231–256, 1998.
- [44] Gonzalo J. Olmo, Diego Rubiera-Garcia, and Aneta Wojnar. Stellar structure models in modified theories of gravity: Lessons and challenges. *Physics Reports*, 876:1–75, 2020. Stellar structure models in modified theories of gravity: Lessons and challenges.
- [45] Madappa Prakash, Manju Prakash, R. Venugopalan, and G. M. Welke. Review of the properties of nuclei in the temperature-density plane. *Phys.Rev. C*, 58:1804–1828, 1998.
- [46] James M. Lattimer, Madappa Prakash, Dieter Masak, and Amos Yahil. Rapidly Rotating Pulsars and the Equation of State. , 355:241, May 1990.
- [47] M. Baldo, I. Bombaci, and G. F. Burgio. Microscopic nuclear equation of state with three-body forces and neutron star structure, 1997.17
- [48] J. W. T. Hessels, S. M. Ransom, I. H. Stairs, P. C. C. Freire, V. M. Kaspi, and F. Camilo. A Radio Pulsar Spinning at 716 Hz. In *American Astronomical Society Meeting Abstracts #207*, volume 207 of *American Astronomical Society Meeting Abstracts*, page 209.07, June 2006.
- [49] W Q Su, J L Han, P F Wang, J P Yuan, Chen Wang, D J Zhou, Tao Wang, Yi Yan, W C Jing, Z L Yang, N N Cai, Xue Chen, Jun Xu, Lang Xie, H G Wang, R X Xu, and X P You. The fast galactic plane pulsar snapshot survey iii. timing results of 30 newly discovered pulsars. *Monthly Notices of the Royal Astronomical Society*, 526(2):2645–2656, September 2023.
- [50] Yi Yan, P F Wang, and J L Han. Distinct polarization properties for two emission states of four pulsars. *Monthly Notices of the Royal Astronomical Society*, 523(4):5207–5219, June 2023.

- [51] P. F. Wang, J. L. Han, J. Xu, C. Wang, Y. Yan, W. C. Jing, W. Q. Su, D. J. Zhou, and T. Wang. Fast pulsar database: I. polarization profiles of 682 pulsars, 2023.
- [52] J. L. Han, Chen Wang, P. F. Wang, Tao Wang, D. J. Zhou, Jing-Hai Sun, Yi Yan, Wei-Qi Su, Wei-Cong Jing, Xue Chen, X. Y. Gao, Li-Gang Hou, Jun Xu, K. J. Lee, Na Wang, Peng Jiang, Ren-Xin Xu, Jun Yan, Heng-Qian Gan, Xin Guan, Wen-Jun Huang, Jin-Chen Jiang, Hui Li, Yun-Peng Men, Chun Sun, Bo-Jun Wang, H. G. Wang, Shuang-Qiang Wang, Jin-Tao Xie, Heng Xu, Rui Yao, Xiao-Peng You, D. J. Yu, Jian-Ping Yuan, Rai Yuen, Chun-Feng Zhang, and Yan Zhu. The fast galactic plane pulsar snapshot survey: I. project design and pulsar discoveries. *Research in Astronomy and Astrophysics*, 21(5):107, jun 2021.
- [53] Zhang, C. M., Wang, J., Zhao, Y. H., Yin, H. X., Song, L. M., Menezes, D. P., Wickramasinghe, D. T., Ferrario, L., and Chardonnet, P. Study of measured pulsar masses and their possible conclusions. *AA*, 527:A83, 2011.

**Project NeoWave: A Comparative Analysis of Conservatism and Gender Dynamics
Impacting Women and Families in the USA and India Since 2010, with Policy
Recommendations By Shivansh Gupta**

Abstract

This study examines the impact of conservative ideologies on gender dynamics in the USA and India since 2010. Using qualitative and quantitative methods, it explores how conservative movements, policies, and cultural norms influence gender roles and expectations. An extensive literature review and policy analysis reveal how conservatism—defined in both countries with its emphasis on profit, output, and traditional gender roles—has shaped systemic oppression and social structures.

The research traces conservatism's historical evolution, highlighting factors like religion, post-colonialism, male dominance, economics, and cultural influences. It discusses significant events affecting gender equality, focusing on the BJP's policies in India and the Republican Party's influence during the Trump administration in the USA. The study evaluates these parties' impacts on women's representation, economic opportunities, and social norms.

Central to the analysis is Project NeoWave, the fifth wave of feminism, which advocates nuanced rights for women, systemic reforms, and new approaches to gender equity. The study compares conservatism's political, social, and economic effects in both countries, examining employment, wage gaps, societal expectations, and the role of traditional values. It also evaluates policies addressing gender and minority development, proposing recommendations that integrate Project NeoWave's principles with conservative frameworks.

By advancing the discourse on conservatism and gender dynamics, the study offers insights for policymakers and advocates, aiming to foster an equitable society through tailored strategies that balance progressive and conservative values.

Keywords

Conservative ideologies, Gender dynamics, Women and families, USA, India, Socio-political landscapes, Cultural norms, BJP policies, Trump administration, Project NeoWave, Fifth wave of feminism, Radical feminism, Comparative analysis, Policy recommendations, Gender minority development

I. INTRODUCTION

DEFINITION OF CONSERVATISM

Conservatism is often viewed primarily as a political ideology, but it is essential to understand it in a broader, more abstract context for this paper. Conservatism should be seen as a system of beliefs deeply intertwined with cultural and religious norms rather than confined solely to the political arena. This distinction is crucial for understanding the social dynamics of women

under conservative ideologies, as it extends beyond political conservatism to encompass the maintenance of traditional power dynamics through culturally ingrained norms and institutions^[1].

About women and current gender dynamics, conservatism typically exerts a restrictive influence on legislative measures that could benefit women, often prioritising cultural preservation over utilitarian outcomes. For instance, the effective implementation of anti-violence against women (VAW) legislation requires complex policy frameworks that frequently clash with conservative values^[2]. Based on a critical analysis of multiple sources, conservatism as it pertains to women can be understood in three key dimensions:

1. **Traditional Values and Gender Roles:** Conservatism often dictates strict societal norms that women must adhere to, reinforcing predefined gender roles. Bangladesh is an example of this, where cultural norms heavily influence women's roles in society^[3].
2. **Conservatism as a Purposeful Way of Life:** For many women, conservatism is not merely a social regulation but a meaningful way of living. Domestic work, for instance, is often perceived not as forced labour but as a choice shaped by upbringing in a conservative household^[4].
3. **Conservatism and Familial Structure:** Conservatism is deeply rooted in family structures, often forming the core of one's identity and self-worth. This familial emphasis strengthens conservative values, embedding them within individuals' personal identities^[5].

This paper delves into the impact of conservative ideologies on gender dynamics in the USA and India, analysing socio-political landscapes, cultural norms, and policies. It also explores the emergence of Project NeoWave, the fifth wave of feminism, through a comparative analysis and proposes policy recommendations to support gender and minority development.

IMPORTANCE OF STUDYING CONSERVATISM'S IMPACT ON GENDER DYNAMICS

Understanding conservatism's impact on gender dynamics is crucial for addressing systemic gender inequalities that permeate various aspects of society, including social, emotional, work, economic, and political spheres. By examining how conservative ideologies shape gender roles and expectations, we can better identify the root causes of these inequalities and develop targeted solutions.

1. **Addressing Systemic Gender Inequality:** Analyzing conservatism's role in perpetuating gender inequality is essential for tackling issues that affect a nation's social, emotional, work, economic, and political conditions. This understanding is vital for

creating effective interventions addressing gender-based disparities' foundational causes^[7].

2. **Supporting Marginalized and Vulnerable Women:** The study is significant for aiding the most marginalised and vulnerable women who face systemic oppression, lack of autonomy, historical marginalisation, and limited opportunities for improvement. By highlighting the impact of conservative values on these groups, policymakers can craft more inclusive and supportive strategies to empower these women.

3. **Informing Policy Development:** Investigating conservatism's influence on gender dynamics provides critical insights for future policy-making. By addressing one of the most fundamental causes of gender discrimination, this research can help create nuanced policies that more effectively tackle gender issues, ensuring that legislative measures are well-informed and impactful.

4. **Enhancing Understanding of Human Dynamics:** Studying conservatism's impact on gender roles offers a deeper understanding of human dynamics, allowing for developing policies more tailored to individual needs. This approach also considers the influence of male-dominated values and their effects on various stakeholders, promoting a more equitable and inclusive societal framework.

This multifaceted exploration of conservatism's impact on gender dynamics is crucial for addressing current gender inequalities and laying the groundwork for a more just and equitable future^[8].

OVERVIEW OF INDIA AND USA'S SOCIO-POLITICAL LANDSCAPES

India's socio-political landscape is characterised by its complex and diverse democracy, with a multi-party system reflecting its vast cultural, linguistic, and regional diversity. The Bharatiya Janata Party (BJP), led by Prime Minister Narendra Modi, has dominated Indian politics recently, particularly since 2014. Focusing on Hindu nationalism, economic reforms, and significant infrastructural development has marked the BJP's governance. However, the party's policies have also sparked controversies, especially concerning issues like religious intolerance, the treatment of minorities, and challenges to democratic institutions. India's political environment continues to evolve, with debates over secularism, federalism, and social justice playing crucial roles in shaping the nation's future^[9].

As of August 2024, the United States is entrenched in a highly polarised socio-political landscape, with the upcoming presidential election amplifying existing divisions. Economic concerns, particularly inflation and the cost of living, are paramount, influencing voter sentiment and policy debates across the political spectrum. This anxiety is compounded by partisan divides,

with Democrats and Republicans sharply opposed on critical issues such as immigration, healthcare, and gun control.

The "culture wars" continue to escalate, with reproductive rights, LGBTQ+ issues, and racial equality at the forefront. In this context, figures like JD Vance have fueled the polarisation further by attacking Kamala Harris's personal life, labelling her a "cat lady" for not having children, thus invoking traditional conservative critiques of women in power. Vance's journey from poverty in Kentucky to Yale Law School—immortalised in his memoir *Hillbilly Elegy* (now a Netflix film)—has also resonated with specific segments of the electorate, who view it as a rags-to-riches story that aligns with the values of self-determination and resilience in a divided America. However, this narrative is contentious, as critiques of his stance on reproductive choices, such as IVF, continue to surface, reflecting deeper ideological battles over family values.

Moreover, the legacy of the Trump administration looms large as populist sentiments and debates over national identity persist, shaping the broader political narrative. These dynamics influence the immediate political climate and hold far-reaching implications for the future of American governance and social cohesion^[10].

II. UNDERSTANDING CONSERVATISM

HISTORICAL CONTEXT OF CONSERVATISM IN INDIA AND THE US

Conservatism in India and the United States has evolved under different historical conditions, yet both share a strong foundation in traditional values and national identity.

In India, conservatism is deeply intertwined with post-colonial nationalism and Hindu religious values. The Bharatiya Janata Party (BJP) rose in 2014 under Narendra Modi and has seen a resurgence of Hindu nationalism, influencing both political and cultural norms. The BJP's governance emphasises traditional family structures and economic conservatism, often at the expense of progressive gender reforms^[11].

In contrast, conservatism in the United States, particularly under the Republican Party, has emphasised limited government, free-market economics, and traditional Judeo-Christian values. The Trump administration saw a resurgence of populist conservatism, with significant opposition to progressive social movements such as LGBTQ+ rights and gender equality.

FUNDAMENTAL PRINCIPLES AND BELIEFS OF CONSERVATIVE IDEOLOGY

The core beliefs of conservatism in both nations include:

- A commitment to traditional family structures and gender roles.
- Opposition to large-scale government intervention, particularly in economic affairs.
- A strong national identity is often tied to religious values.

REASONS FOR THE PERSISTENCE OF CONSERVATISM

Conservatism persists due to its deep cultural and religious roots, which offer stability amidst rapid social changes. In India, the persistence of caste and religious identity reinforces conservative ideologies. In the United States, economic insecurity and the fear of cultural displacement drive many toward conservative political movements, especially among rural and working-class populations.

III. CONSERVATISM AND GENDER DYNAMICS

THEORETICAL FRAMEWORK: GENDER ROLES AND EXPECTATIONS UNDER CONSERVATIVE IDEOLOGIES

Conservative ideologies often reinforce traditional gender roles, where women are expected to fulfil domestic duties and uphold family values. This framework views the family as the central unit of society, with women playing a subordinate but essential role.

ARGUMENTS FOR AND AGAINST CONSERVATISM IN GENDER DYNAMICS

Proponents of conservatism argue that traditional gender roles provide stability and ensure the well-being of families and children. They believe that conservatism preserves cultural heritage and moral values.

Opponents argue that these ideologies perpetuate gender inequality by restricting women's rights and opportunities, particularly in the workforce. Critics highlight the economic disadvantages women face under conservative policies, including wage gaps and limited career advancement.

IV. CONSERVATISM IN INDIA

POLITICAL INFLUENCE: BJP'S POLICIES SINCE 2014

Since coming to power in 2014, the BJP has promoted policies that reinforce traditional gender roles. While some economic initiatives have sought to empower women, the focus on Hindu nationalism and cultural conservatism often marginalises WGM individuals. Policies such as the Triple Talaq ban have been framed as empowering women but are criticised for their selective application of gender justice.

ECONOMIC IMPACT: OPPORTUNITIES AND CHALLENGES FOR WOMEN AND WGM INDIVIDUALS

The economic impact of conservatism in India is mixed. While women have opportunities in sectors such as education and healthcare, conservative cultural expectations limit broader participation in the workforce. Women's labour force participation remains one of the lowest among G20 nations, and the wage gap continues to widen, particularly in rural areas.

SOCIAL AND CULTURAL EXPECTATIONS: TRADITIONAL GENDER ROLES AND FAMILY STRUCTURES

Conservatism in India is closely tied to family structures, where women are expected to maintain traditional roles as caretakers and homemakers. These roles are often reinforced

through religious teachings and societal expectations, limiting women's autonomy and professional growth.

POLICY AND MITIGATION EFFORTS FOR GENDER AND MINORITY DEVELOPMENT

Policies such as the Beti Bachao Beti Padhao (Save the Girl Child, Educate the Girl Child) initiative aim to improve gender equity. However, deeply ingrained cultural norms often undermine these efforts. While some successes have been made, such as enhanced female literacy rates, broader structural changes remain slow.

V. CONSERVATISM IN THE UNITED STATES

POLITICAL INFLUENCE: THE REPUBLICAN PARTY AND THE TRUMP ADMINISTRATION

Under the Trump administration, conservatism in the United States saw a strong resurgence, particularly in opposition to progressive policies regarding gender and minority rights. Trump's conservative Supreme Court appointments and policies restricting reproductive rights and LGBTQ+ protections have had lasting impacts on gender dynamics. The influence of figures such as JD Vance, who used traditional family values as a platform to critique progressive leaders like Kamala Harris, exemplifies the persistence of conservative ideologies in the US.

ECONOMIC IMPACT: OPPORTUNITIES AND CHALLENGES FOR WOMEN AND WGM INDIVIDUALS

While some conservative policies promote individual entrepreneurship and deregulation, they often disadvantage women and WGM individuals, particularly in terms of wage equality and employment opportunities. The rollback of affirmative action and Title IX protections under the Trump administration has disproportionately affected gender and racial minorities, limiting opportunities for equitable hiring and career advancement.

SOCIAL AND CULTURAL EXPECTATIONS: GENDER ROLES AND SOCIETAL NORMS

Conservative values in the US emphasise traditional gender roles, often positioning women as primary caregivers and men as breadwinners. This cultural framework limits the social mobility of women and reinforces gender-based inequalities, particularly in rural and religious communities^[12].

POLICY AND MITIGATION EFFORTS FOR GENDER AND MINORITY DEVELOPMENT

Under the Biden administration, efforts have been made to reverse some of the conservative policies that disproportionately affected gender minorities. Title IX reforms and initiatives promoting diversity in hiring have sought to address these inequalities. However, the Supreme Court's ruling on affirmative action represents a significant setback for gender and racial equality, highlighting the ongoing struggle between progressive and conservative forces in US policy-making^[13].

VI. PROJECT NEOWAVE: THE FIFTH WAVE OF FEMINISM

DEFINITION AND EMERGENCE OF PROJECT NEOWAVE

Project Neowave represents a conceptual "Fifth Wave of Feminism," designed to address current limitations within the feminist movement, drawing on elements of conservatism while maintaining the core goal of empowering women. It proposes to blend conservative values and progressive feminist ideals, creating a more adaptable framework that can better appeal to a broader audience, including those who align with conservative ideologies. To understand where Project Neowave fits into the broader feminist movement, it's essential to first outline the four historical waves of feminism. The First Wave focused on voting, the Second Wave on reproductive rights and workplace equality, the Third Wave introduced more intersectionality, and the Fourth Wave used digital tools and social movements to combat sexual harassment and other forms of systemic inequality^[14].

Project Neowave builds on these foundations by proposing a new regulatory approach for policy change. It does not seek to replace previous feminist efforts but rather to address persistent issues, particularly conservatism extremism and radical feminism, by integrating practical conservative ideals into feminist advocacy.

KEY CHARACTERISTICS OF THE NEW WAVE OF FEMINISM

1. **Conservatism Extremism:** Project Neowave aims to counter the use of conservative ideology to justify barriers for women, such as systemic discrimination and domestic violence. It emphasises using conservative frameworks to uplift women's rights rather than hinder them.
2. **Radical Feminism:** While acknowledging the contributions of radical feminism, Project Neowave critiques its sometimes overly aggressive approach, which can alienate potential allies and discredit the movement. For instance, where radical feminism equates all forms of chivalry with patriarchy might be seen as counterproductive.
3. **Gender Discrimination and Patriarchy:** The project recognises the continued existence of gendered discrimination and systemic patriarchy but offers a pathway for utilising conservative values, like family integrity and responsibility, to promote gender equality in a culturally sensitive manner.

Project Neowave's innovative approach lies in its strategic use of social movements and the current social media mechanisms, including cancel culture, to push for legislative changes. It offers an opportunity to build a bridge between conservative strongholds and the feminist movement by demonstrating that these ideologies can coexist and support each other when properly aligned.

EXAMPLES OF RADICAL FEMINISM AND ITS INFLUENCE

Historically, radical feminism emerged as a critique of the societal structures that uphold patriarchy. However, in recent years, some radical feminist perspectives have been criticised for their exclusionary or extreme views. For example, rejecting practices such as chivalry outright as oppressive might alienate those who believe these acts come from positive intentions, thus losing support from potential male allies and even women who may not identify with such extreme positions^[15].

Project Neowave seeks to incorporate these critiques by advocating for balance. It aims to temper some of radical feminism's more divisive tendencies to create a movement that is both inclusive and effective without losing sight of its core objectives.

EFFORTS WITHIN PROJECT NEOWAVE TO ADDRESS GENDER MINORITY DEVELOPMENT

Project Neowave is distinct in its application of feminist principles to support Women and Gender Minorities (WGM):

1. **Curbing Liberal Radicalism:** It promotes the idea that extreme liberalism, especially when unmoderated, can alienate gender minorities and even those within the feminist movement who advocate for incremental change.
2. **Integrating Conservative Learnings:** Project Neowave draws on conservative lessons, such as the value of family unity and community support, to offer a more inclusive environment for WGM individuals.
3. **Advocacy Through Social Movements:** It acknowledges that while court appeals and government challenges are valid, social movements—especially in the age of social media—are the most effective way to bring about legislative change for WGMs and broader society.

CRITIQUES AND POTENTIAL HARMS OF RADICAL FEMINISM

1. **Extremism Alienates Potential Allies:** Radical feminism's confrontational tactics can sometimes overshadow the movement's core goals, alienating moderates and reinforcing negative stereotypes. This could result in a lack of mainstream support.
2. **Discouraging Men's Participation:** When feminism is perceived as being hostile toward men, it can discourage male allies from engaging in the movement, which is counterproductive to gender equality goals.
3. **Exclusion of Vulnerable Women:** Radical feminism can sometimes be inaccessible to the most vulnerable women, particularly those in conservative societies or

lower socio-economic statuses. Project Neowave, however, aims to build bridges rather than walls, ensuring the movement is open to all women.

DISCUSSION ON THE IDEAL MODEL MOVEMENT THAT

INTEGRATES CONSERVATIVE FACTORS TO SUPPORT THE MOST VULNERABLE INDIVIDUALS EFFECTIVELY

New Domesticity is one of Project Neowave's strategies to integrate conservative ideals into the feminist movement. It seeks to reframe traditional domestic roles—such as caregiving and homemaking—not as limitations on women but as empowering choices. This concept is precious because it appeals to conservative men and women by reinforcing values they already uphold while allowing space for progressive empowerment. The goal is not to insist that women should only be homemakers but to validate these roles for those who choose them. This validation builds a bridge between conservative and feminist ideologies, showing that they can coexist in a way that supports the most vulnerable individuals.

VII. COMPARATIVE ANALYSIS

Conservatism as a political and social ideology has had profound and lasting effects on gender dynamics, both in the United States and India. While the specific cultural and political environments of each country differ, there are significant similarities in how conservatism shapes gender roles, political representation, and economic opportunities for women and gender minorities (WGM). However, the differences are also crucial in understanding conservatism's diverse challenges and consequences for gender equality in these nations.

SIMILARITIES IN CONSERVATISM'S IMPACT ON GENDER DYNAMICS

ECONOMIC: EMPLOYMENT OPPORTUNITIES AND WAGE GAPS

In both India and the United States, conservatism has contributed to persistent wage gaps and limited employment opportunities for women and gender minorities. This limitation is often tied to the traditional conservative belief that men are the primary breadwinners while women are better suited to roles within the household. As a result, conservative policies in both nations tend to resist legislative efforts to bridge wage disparities and provide equal employment opportunities for WGM individuals.

In India, women's participation in the workforce is significantly lower, particularly in rural areas, where conservative norms are more rigidly enforced. For instance, the cultural expectation that women prioritise caregiving roles often excludes them from the formal economy. According to Kabeer and Mahmud, women in India face unique challenges within the informal economy, where they are more vulnerable to wage discrimination and lack labour protections.

Similarly, in the United States, while the labour force participation rate for women is higher, wage gaps persist, especially in conservative-dominated regions. Conservative opposition

to progressive economic policies, such as gender-based affirmative action and equal pay legislation, has hampered addressing these inequalities.

The Trump administration in the U.S. exemplifies how conservative economic policies can disadvantage women and gender minorities. By rolling back Title IX protections and limiting workplace equality initiatives, the administration exacerbated existing disparities in pay and employment opportunities for women and WGM individuals^[16].

Similarly, in India, while economic initiatives such as Beti Bachao Beti Padhao have sought to empower women, the conservative focus on maintaining traditional family structures has often undermined these efforts.

POLITICAL: REPRESENTATION AND PARTICIPATION OF WOMEN AND WGM INDIVIDUALS

Conservatism in both the U.S. and India has led to the underrepresentation of women and gender minorities in political leadership roles. Political conservatism often advocates maintaining traditional power structures, where men dominate leadership positions and decision-making processes. The resistance to gender quotas and other affirmative action policies is a hallmark of conservative governance in both countries, resulting in fewer women in key political positions.

In India, the Bharatiya Janata Party (BJP), since coming to power in 2014, has reinforced traditional gender roles through policies that marginalise WGM individuals. Despite some reforms, such as the Triple Talaq ban, which was touted as empowering Muslim women, critics argue that these initiatives have been selective in applying gender justice and have failed to address the broader issue of gender equality.

The lack of representation for women in political leadership positions remains a critical issue, with conservative elements within the BJP resisting efforts to introduce gender quotas or other reforms that would ensure a more equitable distribution of political power.

Similarly, in the United States, the conservative influence under the Trump administration saw a decline in efforts to promote gender equality in political representation. By appointing conservative judges to the Supreme Court and rolling back key policies related to reproductive rights and gender protections, the administration reinforced traditional power structures that limit the participation of women and WGM individuals in politics. A study published by the Council on Foreign Relations highlights how conservative governance often leads to lower levels of women's representation in decision-making roles, further entrenching gender disparities^[17].

SOCIAL: SOCIETAL EXPECTATIONS AND TRADITIONAL GENDER ROLES

In both nations, conservatism emphasises traditional gender roles, where women are primarily seen as caregivers and men as breadwinners. These roles are reinforced through cultural norms, religious teachings, and conservative social policies. In India, Hindu conservatism strongly supports the idea that women's primary role is within the home, taking care of family and children. This expectation is prevalent, especially in rural areas, where women's autonomy and professional opportunities are severely limited by societal norms.

Similarly, in the United States, Christian Evangelical values play a significant role in shaping conservative views on gender roles. Conservative rhetoric in the U.S. often emphasises the importance of the nuclear family, with women expected to prioritise motherhood and homemaking over professional aspirations. This cultural framework limits social mobility for women and perpetuates gender-based inequalities, particularly in more rural and religious communities.

CULTURAL: INFLUENCE OF RELIGION AND TRADITIONAL VALUES

India and the United States have deeply rooted religious and cultural traditions that shape their conservative values. In India, Hindu conservatism promotes a cultural framework that emphasises the importance of family and traditional gender roles. This religious influence extends into political and economic life, where women are often expected to conform to societal norms that prioritise family responsibilities over professional or political ambitions.

In the United States, Christian conservatism similarly upholds traditional family structures and gender roles. Evangelical Christianity plays a central role in shaping conservative values around gender, particularly in opposition to reproductive rights, same-sex marriage, and gender identity protections. The emphasis on family and traditional roles significantly impacts how conservative policies are crafted and implemented in both nations.

DIFFERENCES IN CONSERVATISM'S IMPACT ON GENDER DYNAMICS

ECONOMIC: VARYING LEVELS OF ACCESS AND OPPORTUNITY

While both countries experience wage gaps and employment disparities under conservative governance, the specific challenges faced by women and WGM individuals differ significantly. In India, women's access to economic opportunities is severely restricted, particularly in rural areas where conservative cultural values are more rigidly enforced. The informal economy plays a significant role in women's employment in India. Still, it also exposes them to more significant risks of wage exploitation and lacks the protections afforded to workers in the formal economy^[18].

By contrast, in the United States, conservative policies often promote entrepreneurship and deregulation, which can create opportunities for some women. However, these same policies also undermine critical protections, such as Title IX, which is essential for ensuring equal access to employment and education for women and WGM individuals.

POLITICAL: DIFFERENCES IN POLICY APPROACHES AND OUTCOMES

Conservative governance in India and the United States takes different approaches to gender-based political reforms. In India, conservative politics under the BJP is closely tied to Hindu nationalism, which influences how gender policies are framed. For example, the Triple Talaq ban, while framed as empowering Muslim women, has been critiqued for selectively applying gender justice without addressing broader issues of gender inequality.

In contrast, the U.S. has seen a conservative focus on curtailing reproductive rights and restricting gender identity protections, primarily through judicial appointments and legislative changes made during the Trump administration.

SOCIAL: DIVERSE SOCIETAL REACTIONS AND ADAPTATIONS

The social reactions to conservatism also vary significantly between the two countries. In India, traditional family structures remain rigid, and societal norms strongly support conservative values that limit women's autonomy. In contrast, the U.S. experiences diverse societal reactions to conservatism. While rural and religious communities tend to uphold traditional gender norms, urban and progressive regions offer more excellent resistance to conservative policies and social expectations.

CULTURAL: DISTINCT CULTURAL PRACTICES AND BELIEFS

The cultural differences between India and the United States are evident in how conservatism influences gender dynamics. In India, Hindu nationalism plays a central role in shaping conservative gender norms. In contrast, in the United States, Christian conservatism informs debates around reproductive rights, same-sex marriage, and gender identity protections. These cultural influences profoundly impact how gender roles are constructed and maintained in each nation.

VIII. POLICY AND MITIGATION EFFORTS

Significant efforts have been made in both India and the United States to promote gender and minority development through various policies. However, these policies' outcomes and effectiveness vary due to political, social, and economic differences.

EVALUATION OF POLICIES AIMED AT GENDER AND MINORITY DEVELOPMENT IN BOTH COUNTRIES

In India, The Bharatiya Janata Party (BJP) has implemented several policies to promote gender equality, including social welfare schemes like Beti Bachao Beti Padhao (Save the Girl Child, Educate the Girl Child). However, critics argue that while these initiatives have improved education and awareness, they have not fully addressed deep-seated patriarchal norms. India's gender quota law, which reserves seats for women in local governments, has proven to be a successful initiative for improving women's political representation at the grassroots level. Research shows that these quotas have increased women's participation in politics and governance, leading to more policies that cater to female citizens' needs, such as public goods provision and education for girls^[19].

Despite these successes, challenges remain. Traditional gender norms and a lack of political will from male party leaders continue to limit the broader advancement of women in higher political offices. Furthermore, minority women, particularly Muslim women, have been at the forefront of protests against government policies, highlighting the ongoing struggle for gender and minority rights.

In the United States, the Biden administration has introduced the National Gender Strategy, the first-ever comprehensive federal strategy to address gender inequality. This strategy focuses on ten priorities, including improving women's economic security, ensuring access to healthcare (including reproductive services), and eliminating gender-based violence. The approach also emphasises intersectionality, recognising the compounded discrimination faced by racial, sexual, and gender minorities^[20].

While the strategy aims to address systemic gender inequalities, the U.S. has faced setbacks, especially during the Trump administration, which saw rollbacks in policies related to reproductive rights and LGBTQ+ protections. Despite this, initiatives such as Title IX reforms under the current administration are working to ensure equal opportunities in education and employment for women and minorities. The U.S. struggles to close the wage gap and increase political representation for women and minorities. However, there has been notable progress in electing women to Congress and other leadership roles^[21].

SUCCESSFUL INITIATIVES AND THEIR IMPACT

Gender Quotas in India: Implementing gender quotas in local governments has had a positive impact on women's representation, governance responsiveness, and policy outcomes favoring women. However, these reforms have also faced backlash, particularly from male leaders who resist challenges to traditional power structures.

National Gender Strategy in the U.S.: The Biden administration's strategy is notable for its whole-of-government approach, emphasising collaboration across federal, state, and local levels to mainstream gender equality into all aspects of governance and policy. The strategy has the potential to make substantial progress in economic security, healthcare access, and addressing gender-based violence.

RECOMMENDATIONS FOR FUTURE POLICY AND ADVOCACY EFFORTS

GENDER-RESPONSIVE CONDITIONAL GRANTS

- **Problem:** Persistent gender wage gaps and lack of economic opportunities for women and gender minorities.
- **Recommendation:** Introduce Gender-Responsive Conditional Grants that provide direct funding to companies, NGOs, and startups that meet gender-equality benchmarks. This can include transparent hiring practices, equal pay audits, and female representation in leadership. Companies that close the wage gap and promote diversity would receive additional tax benefits or funding, creating financial incentives for inclusivity.
- **Inspired by:** The success of conditional cash transfers in social programs like India's Beti Bachao Beti Padhao, but with a focus on corporate gender parity.

WORK-LIFE INTEGRATION POLICIES

- **Problem:** Traditional gender roles in both countries expect women to manage household responsibilities alongside professional duties, reducing career advancement opportunities.
- **Recommendation:** Implement Work-Life Integration Policies that promote paid caregiving leave for both genders in all sectors. Rather than maternity leave focusing solely on women, this policy would mandate that men take equal caregiving responsibilities. By framing caregiving as a societal duty and including incentives for employers who adopt flexible work arrangements, we can address ingrained social norms limiting women's progress.
- **Inspired by** Sweden's paternity leave policy, which enforces leave for men to ensure shared caregiving responsibilities.

INTERSECTIONAL ADVOCACY HUBS

- **Problem:** Gender policies often fail to address the compounded challenges faced by minority women and gender minorities (e.g., Muslim women in India and LGBTQ+ individuals in the U.S.).
- **Recommendation:** Establish Intersectional Advocacy Hubs as resource centres for minority women and WGM individuals, providing access to legal aid, job training, mental health resources, and financial literacy programs. These hubs would also serve as policy incubators, where experts and marginalised community members work together to develop laws that cater to intersectional challenges. This would ensure that policies are shaped by the lived experiences of those most affected by systemic discrimination.
- **Inspired by** The Biden administration's intersectional approach to gender equity

GENDER EQUALITY INVESTMENT FRAMEWORK (GEIF)

- **Problem:** Systemic underinvestment in women-led enterprises and gender-diverse organisations.
- **Recommendation:** Create a Gender Equality Investment Framework (GEIF), a public-private investment model with funds allocated explicitly to women-led startups and businesses promoting gender equality. This could be coupled with public procurement policies prioritising gender-inclusive companies for government contracts. Additionally, community investment bonds can encourage local participation in the growth of women-led businesses.
- **Inspired by:** Public procurement initiatives in several European countries that reward inclusive companies

INCLUSIVE DIGITAL LITERACY PROGRAMS

- **Problem:** Women and gender minorities face a significant digital divide, particularly in rural areas.
- **Launch Inclusive Digital Literacy Programs** to close the digital divide for women in rural areas and marginalised communities. This initiative would focus on training women and WGM individuals in technical and digital skills, providing them with the tools to participate in modern economies. In partnership with telecom companies, governments could offer subsidised digital devices and internet access in underserved areas, combined with sector-specific training.
- **Inspired by** The World Bank's strategy of increasing access to technology for women and underserved groups

COMMUNITY-LED GENDER VIOLENCE MONITORING UNITS

- **Problem:** Gender-based violence is rampant, with legal systems often failing to provide timely justice, particularly in rural areas.
- **Recommendation:** Establish Community-Led Gender Violence Monitoring Units where local volunteers are trained in gender-based violence (GBV) intervention and legal processes. These units, supported by local governments and NGOs, would ensure that victims receive immediate support, from medical care to legal aid. The units would also collaborate with law enforcement agencies to monitor and report gender-based crimes, ensuring accountability.
- **Inspired by** India's One-Stop Centers for GBV victims, but expanded into community-led initiatives

CULTURAL CHANGE CAMPAIGNS

- **Problem:** Deep-rooted cultural norms that reinforce traditional gender roles and limit the aspirations of women and gender minorities.
- **Recommendation:** Governments and NGOs should launch Cultural Change Campaigns that use media, entertainment, and education to challenge gender stereotypes and promote gender equality. These campaigns could feature role models from diverse backgrounds, showing women in leadership, STEM, and other non-traditional roles. By working with influencers and cultural leaders, these campaigns aim to shift public perceptions of gender roles.
- **Inspired by:** The "This Girl Can" campaign in the UK successfully shifted perceptions about women in sports.

DECENTRALISED GENDER AUDITS

- **Problem:** Gender inequality within companies and institutions often goes unnoticed due to a lack of transparency and accountability.
- **Recommendation:** Mandate Decentralized Gender Audits conducted by independent auditors who evaluate companies' and government institutions' compliance with gender equality goals. These audits would assess areas like wage equality, leadership representation, and anti-discrimination policies. Companies that meet gender parity goals would receive public recognition and tax incentives, while those failing to comply would face penalties.
- **Inspired by:** The gender parity assessments several Scandinavian countries use to ensure workplace equality.

INTEGRATING PROJECT NEOWAVE'S PRINCIPLES WITH CONSERVATIVE VALUES TO CREATE EFFECTIVE POLICIES

Project Neowave proposes a nuanced approach to gender equality, integrating progressive feminist ideals with conservative values to create sustainable policies. Project Neowave can push for policies that respect conservative cultural norms while advancing women's rights by focusing on commonalities, such as the importance of family and community. For example, promoting the idea of new domesticity, which values caregiving roles traditionally held by women while advocating for equal economic opportunities, can appeal to conservative bases in both countries.

Project Neowave can also advocate for gender-sensitive policies that emphasise family integrity and community support, presenting them as beneficial to society. By framing gender equality as a means to strengthen communities and economies rather than a threat to traditional values, Project Neowave could bridge the gap between conservative and feminist ideologies.

In conclusion, while both countries have made progress in promoting gender and minority rights, future efforts must address the remaining structural challenges. By integrating Project Neowave's principles into policy discussions, both India and the United States could find ways to balance progressive reforms with conservative values, ultimately creating more inclusive and sustainable policies.

These creative policy recommendations aim to tackle the root causes of gender inequality by integrating elements from previous successful initiatives while adding new, forward-thinking strategies. They focus on economic empowerment, intersectional inclusion, and systemic change, ensuring that women and gender minorities can thrive in India and the United States. By promoting policies that balance conservative values with progressive goals, Project Neowave seeks to create sustainable and inclusive pathways for gender and minority development.

IX. PERSONAL REVIEW/REFLECTION

Reflecting on this research, I've experienced a profound shift in my understanding of conservatism and its intricate role in shaping gender dynamics. At the start of this journey, I saw conservatism as purely restrictive, a force that sought to confine women within rigid, patriarchal structures. Coming from a conservative family, I have seen this firsthand—traditional gender

roles being reinforced through culture and religion. It often felt like conservatism was synonymous with oppression, a system designed to limit women's opportunities and keep them tethered to societal norms that I had always seen as stifling. My initial perception was that conservatism primarily served the interests of men, especially in maintaining power and control over women. Being a cisgender male, I've had the privilege of navigating through life with relative ease. Still, I have seen how the same cultural norms have affected the women in my family—limiting their access to opportunities that seemed effortlessly available to me.

However, as I delved deeper into the topic, I realised that conservatism is not necessarily a force for harm. It holds a capacity to preserve cultural identity, foster social cohesion, and even uplift women in ways that liberal frameworks sometimes overlook. It was surprising to learn that conservative women are not always passive victims of oppression but can be active participants in preserving traditional values that provide them with a sense of purpose and meaning. There's a strong argument for how conservatism—particularly cultural and religious values—can offer a stabilising force, protecting against radical disruptions that could potentially harm societal cohesion. For instance, Project Neowave, the conceptual model I've developed, aims to harness the positive aspects of conservatism, like family integrity and community values, and merge them with the feminist push for gender equality. I've come to see that conservatism can, in fact, support women's empowerment by offering them a framework that aligns with their cultural or religious beliefs while also advocating for broader rights and opportunities.

I also discovered that radical feminism, which I initially dismissed as a theoretical exaggeration, does indeed exist and can, at times, cause more harm than good. It became clear to me that while radical feminism has its place in challenging entrenched power structures, it can alienate potential allies and perpetuate the notion that feminism is a divisive movement. The term "radical feminism" isn't merely a myth or an exaggeration; in some contexts, it has played a part in making feminism appear overly combative and disconnected from the needs of everyday women. This realisation challenged my biases, as I had always assumed that the more radical the approach, the more effective it was in dismantling oppressive systems. However, through this research, I learned that radicalism—whether conservative or feminist—can backfire. Change is often more sustainable when it incorporates multiple perspectives, including those of conservative women who may not see themselves as oppressed but rather as empowered through their roles in traditional systems.

One of the most valuable lessons I gained was understanding the plight of Women and Gender Minorities (WGM). While I had previously grouped their experiences with that of women in general, I now recognise the distinct challenges WGM individuals face, especially within conservative frameworks. Their inclusion in this study was crucial because they are among the most vulnerable groups affected by gender-based discrimination, often navigating not just patriarchy but also social stigmatisation that stems from their gender identity. This new understanding deepened my sense of empathy for how complex and layered gender inequality can be—especially for those who do not fit into the binary gender norms reinforced by conservative societies.

Reflecting on my initial views, I can see how much my perspective has evolved. I no longer view conservatism as entirely negative but as a complex system with benefits and drawbacks. If applied thoughtfully, it can empower women by reinforcing strong, community-focused values while embracing necessary reforms for gender equality. On the other hand, radical feminism—once something I saw as a progressive necessity—now feels, in certain instances, like a misstep that hinders the broader feminist movement by alienating people who might otherwise support its goals. I’ve learned that new domesticity, for example, doesn’t have to be regressive but could bridge conservative and feminist ideals, encouraging women to thrive in and outside the home. This balanced approach seems to be the way forward.

In summary, I believe the solution is creating systems that integrate the best aspects of conservative stability and progressive change. Project Neowave represents my vision for a new feminist movement that respects cultural values while pushing for the empowerment of women and WGM individuals. As someone passionate about policy and systems change, I now realise the importance of cultural sensitivity in any advocacy work. My recommendations moving forward would be to push for policies that respect the conservative roots of communities while fostering economic, social, and political opportunities for all genders. This means crafting laws that provide equal access to opportunities without disregarding the traditional values that hold meaning for so many people. In doing so, we can create a system where gender equality is not a threat to tradition but a natural evolution.

X. CONCLUSION

SUMMARY OF KEY FINDINGS

The research revealed several key findings regarding conservatism’s impact on gender dynamics in India and the United States. First, conservatism reinforces traditional gender roles in both countries, where women are often seen as primary caregivers, limiting their professional opportunities. In India, this is driven by Hindu nationalism, while in the U.S., Christian Evangelical values uphold similar norms. This cultural conservatism contributes to wage gaps and employment disparities in both countries, with conservative policies often resisting gender-based reforms like affirmative action and Title IX protections. Politically, conservatism has led to the underrepresentation of women and WGM individuals in leadership roles, with policies like India’s Triple Talaq ban and conservative reforms in the U.S. reinforcing gender inequality in both political and social participation.

Religiously rooted conservative values in both countries create barriers to gender equality, as they emphasise maintaining traditional family structures. However, the study also introduced Project Neowave, a conceptual feminist movement that seeks to integrate conservative values with feminist ideals. It proposes using conservative values like family integrity to bridge the gap between traditional beliefs and modern gender equality goals. Ultimately, the research highlights the complexity of conservatism’s role in shaping gender

dynamics and stresses the need for policies that balance traditional values with progressive reforms to achieve a more inclusive and equitable society.

IMPLICATIONS FOR FUTURE RESEARCH AND POLICY-MAKING

The findings of this research highlight the need for future policy efforts to be more nuanced and inclusive, particularly in addressing the unique challenges WGM individuals face. Future research should delve deeper into the intersectionality of gender, race, and class, especially in conservative contexts where traditional norms are deeply entrenched. There is also a need to explore the role of men as allies in gender equality movements, as the current focus on radical feminism sometimes alienates potential male supporters. Policymakers must focus on culturally sensitive solutions that respect traditional values while ensuring equal opportunities for all genders.

FINAL THOUGHTS ON CONSERVATISM'S ROLE IN SHAPING GENDER DYNAMICS

The research underscores the complex role conservatism plays in shaping gender dynamics in both India and the United States. While conservatism often reinforces patriarchal structures, it can also offer pathways for community-driven empowerment if the values it promotes—such as family, stability, and cultural identity—are harnessed to support progressive gender policies. The persistence of conservatism is not solely due to its opposition to change but because it offers many people a sense of purpose, identity, and belonging.

VISION FOR A MODEL MOVEMENT

My vision for a model movement—Project NeoWave—strikes a balance between conservative and feminist ideals. It is a movement that respects tradition while advocating for policies that address modern challenges such as wage gaps, underrepresentation in politics, and gender-based violence. By integrating conservative values like family unity and social responsibility with feminist goals of empowerment and equality, Project NeoWave can appeal to a broader base, including those who may feel alienated by radical feminism. This balance will be vital in crafting sustainable and effective policies that uplift women and WGM individuals, making them active participants in traditional and progressive spaces.

In summary, by acknowledging the merits of conservatism while advancing feminist ideals, we can create a more inclusive and equitable society that benefits everyone, regardless of gender or cultural background. This research provides the groundwork for future policy efforts and advocacy strategies that are innovative and deeply attuned to the values and beliefs that shape our societies.

XI. ACKNOWLEDGEMENTS

I would like to express my gratitude to Ms. Meaghan Blight (President, Wesleyan College) for extending her help in carrying out, editing, improving and consolidating this research.

Works Cited

- Stanford Encyclopedia of Philosophy. "Conservatism." Stanford Encyclopedia of Philosophy, plato.stanford.edu/entries/conservatism/. Accessed 4 Aug. 2024.
- Htun, Mala, and S. Laurel Weldon. "The Civic Origins of Progressive Policy Change: Combating Violence against Women in Global Perspective, 1975-2005." *Comparative Political Studies*, vol. 54, no. 12, 2021, pp. 2051–2081, <https://journals.sagepub.com/doi/full/10.1177/00104140211024313>. Accessed 4 Aug. 2024.
- Kabeer, Naila, and Simeen Mahmud. "Mainstreaming Gender in Social Protection for the Informal Economy." *World Development*, vol. 103, 2018, pp. 230-240, <https://www.sciencedirect.com/science/article/pii/S0305750X18302080>. Accessed 4 Aug. 2024.
- Hymowitz, Kay Parker. "Why Conservative Women Report Being the Happiest—and How You Can Be Too." American Enterprise Institute, www.aei.org/op-eds/why-conservative-women-report-being-the-happiest-and-how-you-can-be-too/. Accessed 4 Aug. 2024.
- Hallowell, A. J. "Conservatives and the Family." The Russell Kirk Center for Cultural Renewal, kirkcenter.org/kirk-essay-conservatives-and-the-family/. Accessed 4 Aug. 2024.
- Cambridge Core. "Gender Politics and Conservatism: The View from the British Conservative Party Grassroots." *Government and Opposition*, www.cambridge.org/core/journals/government-and-opposition/article/gender-politics-and-conservatism-the-view-from-the-british-conservative-party-grassroots/110FE2674F196BB84C5424F2C88F2444. Accessed 4 Aug. 2024.
- Oxford Academic. "Gender Equality and Conservatism." Oxford University Press, academic.oup.com/edited-volume/28063/chapter/212060501. Accessed 4 Aug. 2024.
- BBC. "India Election 2024: Why it Matters for the World." BBC News, 12 Apr. 2024, www.bbc.com/news/world-asia-india-65047658. Accessed 4 Aug. 2024.
- Pew Research Center. "Top Problems Facing the US: Inflation, Affordable Health Care, and Partisan Cooperation." Pew Research Center, 2024, www.pewresearch.org/fact-tank/2024/08/01/top-problems-facing-the-us-inflation-affordable-health-care-partisan-cooperation/. Accessed 4 Aug. 2024.
- Brookings Institution. "Geopolitical & Economic Outlook 2024: Democracy and the Splintering." Brookings Institution, www.brookings.edu/research/geopolitical-economic-outlook-2024-democracy-and-the-splintering/. Accessed 4 Aug. 2024.
- Evans, Jonathan, et al. "Gender Roles in the Family." Pew Research Center's Religion & Public Life Project, 2 Mar. 2022, www.pewresearch.org/religion/2022/03/02/gender-roles-in-the-family/. Accessed 4 Aug. 2024.

- Pew Research Center. "Public Views on Changing Gender Roles." Pew Research Center, 2023, www.pewresearch.org/fact-tank/2023/05/25/public-views-on-changing-gender-roles/. Accessed 21 Oct. 2024.
- Brookings Institution. "The Biden-Harris Administration's Gender Strategies and Policies: Strengths, Challenges, and Opportunities." Brookings, 2024, <https://www.brookings.edu/articles/the-biden-harris-administrations-gender-strategies-and-policies-strengths-challenges-and-opportunities/>. Accessed 21 Oct. 2024.
- Pew Research Center. "A Century After Women Gained the Right To Vote, Majority of Americans See Work To Do on Gender Equality." Pew Research Center, 2020, www.pewresearch.org/fact-tank/2020/07/07/a-century-after-women-gained-the-right-to-vote-majority-of-americans-see-work-to-do-on-gender-equality/. Accessed 21 Oct. 2024.
- Pew Research Center. "How the Political Typology Groups View Major Issues." Pew Research Center, 9 Nov. 2021, www.pewresearch.org/politics/2021/11/09/how-the-political-typology-groups-view-major-issues/. Accessed 21 Oct. 2024.
- "Catalyzing Gender Equality: A Comprehensive Evaluation of Title IX and Conservative Economic Policies." Cambridge Core, 2023, <https://www.cambridge.org/core/journals/government-and-opposition>. Accessed 1 Oct. 2024.
- Council on Foreign Relations. "The Fight for Gender Parity." Council on Foreign Relations, 2023, www.cfr.org/blog/fight-gender-parity. Accessed 21 Oct. 2024.
- Williams, Colin, and Ardiana Gashi. "Evaluating the Wage Differential between the Formal and Informal Economy: A Gender Perspective." *Journal of Economic Studies*, vol. 49, no. 4, 2021, pp. 735-750, Emerald Publishing Limited, <https://doi.org/10.1108/JES-01-2021-0019>. Accessed 21 Oct. 2024.
- Brechenmacher, Saskia. "India's New Gender Quota Law Is a Win for Women—Mostly." Carnegie Endowment for International Peace, 26 Sep. 2023, www.carnegieendowment.org/indias-new-gender-quota-law-is-a-win-for-women-mostly/. Accessed 21 Oct. 2024.
- The White House. National Strategy on Gender Equity and Equality. The White House, Oct. 2021, www.whitehouse.gov/wp-content/uploads/2021/10/National-Strategy-on-Gender-Equity-and-Equality.pdf. Accessed 21 Oct. 2024.
- World Bank. World Bank Group Gender Strategy 2024–2030: Accelerate Gender Equality to End Poverty on a Livable Planet: World Bank Group, 2024, <https://blogs.worldbank.org/voices/raising-world-bank-groups-ambition-gender-equality>. Accessed 21 Oct. 2024.

The Microchip Revolution: Jack Kilby's Invention and the Past, Present and Future of Technology By Alexander Recce

Abstract

Since the invention of the transistor in 1947, innovations in electronic circuitry have paved the way for modern technology from spaceships to smartphones. The miniaturization of microchips in particular has enabled the rapid development of computer technology that we depend on daily in personal and professional settings. In 1958 two different engineers: Jack Kilby at Texas Instruments (TI) and Robert Noyce at Fairchild Semiconductor made complementary breakthroughs that would lead to the development of the integrated circuit, kicking off the “chip wars” which resulted in the vast array of modern electronics. This historical case study sheds light on the impact of individuals, competition, and the professional and political landscape in the history of computing technology. In particular, we argue that unexpected interactions between multiple minds operating independently are responsible for technological innovation. Finally, we suggest implications from this research for the future of computing.

Keywords

Electronics, History of technology, Jack Kilby, microchips

Introduction

The year 1958 was a pivotal moment in the history of technology. Following World War II, the United States witnessed a surge in economic growth, propelled by a booming population, immense government spending on research and development, and mass consumerism. This decade saw a 37% increase in the US economy⁹⁵ and the emergence of the baby boomers, the largest population cohort in the country's history⁹⁶.

Amidst this backdrop, the transistor had just been invented ten years earlier at Bell Labs by William Shockley, John Bardeen and Walter Brattain⁹⁷. This invention offered a more reliable, efficient and smaller alternative to the vacuum tubes used in electronic devices. Transistors rapidly became integral components in radios, televisions and early computers. Electronic devices that used transistors instead of vacuum tubes were called solid state. There was a lot of excitement about what was possible with solid state technology, but the cost and complexity of the circuitry required to scale up the number of components in electronic devices was prohibitive. This problem was called the interconnection problem⁹⁸ and the “tyranny of

⁹⁵ Halberstam, David. *The fifties* p.10

⁹⁶ Halberstam, David. p. 134

⁹⁷ Reid, Thomas *The chip* p. 11

⁹⁸ As a concrete example, consider a modern personal computer with one gigabyte of memory. This is eight billion binary digits, and each digit requires about five circuit components. Without an IC these 40 billion components would have to be connected to each other in exactly the right way using individual wires. In order to connect the wires the components have to be larger and wiring takes up a lot of space.

numbers.”⁹⁹

During this period the geopolitical landscape also underwent significant transformations. The Cold War and the nuclear arms race were increasing in intensity¹⁰⁰. The Korean War, lasting from June 1950 to July 1953, marked a period of heightened global tensions¹⁰¹. Further escalating the stakes, in October 1957, the Soviet Union launched Sputnik-1, the first artificial Earth satellite¹⁰². This historic event ignited a fierce competition in space exploration and also led to the establishment of NASA in July 1958¹⁰³

All of these geopolitical tensions accelerated the development of computers and technology. In 1958, International Business Machines (IBM) unveiled the 7070 Data Processing System, the first commercially available computer made with transistors¹⁰⁴. The number of transistors is not available, but the model 7090, released a year later, contained 50,000 transistors¹⁰⁵, occupied eight large refrigerator sized cabinets, and the inflation-adjusted price was twenty-nine million dollars in 2024 currency.

Jack Kilby

Born in Jefferson City, Missouri in 1923, Jack Kilby had an active interest in fixing things and electronics from a young age¹⁰⁶. His dad was employed at power companies as an electrical engineer, and he would travel with his father to fix small generators and transformers. These experiences crystalized his view of inventing improvements to solve problems, and a focus on making things simpler and less expensive. He said, “The way that my dad always liked to put it was that an engineer could find a way to do for one dollar what everybody else could do for two.”¹⁰⁷

Jack Kilby began his academic training with an electrical engineering degree from the University of Illinois at Urbana-Champaign¹⁰⁸. Prior to changing jobs to work at TI, Kilby worked eleven years at Centralab in Milwaukee, where he gained valuable experience in developing circuit boards and other components. At the same time, he studied for a master’s degree in electrical engineering at the University of Wisconsin.

He had a deep interest in understanding the physics of how electric devices work, but he never saw himself as a scientist. In his words: “There is a pretty key difference. A scientist is motivated by knowledge, he basically wants to explain something. An engineer’s drive is to solve problems, to make something work.”¹⁰⁹

⁹⁹ Reid, Thomas p. 13

¹⁰⁰ Halberstam, David. *The fifties* p. 15

¹⁰¹ Halberstam, David. p. 73

¹⁰² Halberstam, David. p. 624

¹⁰³ Halberstam, David. p. 628

¹⁰⁴ Bashe, Charles et. al. *IBM's early computers*. p. 446

¹⁰⁵ Bashe, Charles et. al. p. 448

¹⁰⁶ Eckstein, Peter. *Jack Kilby (1923-2005)* p. 90

¹⁰⁷ Reid, Thomas *The chip* p. 67

¹⁰⁸ Reid, Thomas p. 68

¹⁰⁹ Reid, Thomas *The chip* p. 64

Centralab's earlier involvement in World War II's technological advancements set the stage for Kilby's innovations. The company played a crucial role in developing a printed circuit board method for a radio proximity fuse used in mortar shells. This method, which entailed arranging conductive lines on an insulating base to form a 'printed' interconnection design for components, laid the groundwork for future miniaturization¹¹⁰. Additionally, Centralab's work on printed circuit audio amplifier circuits underscored its pioneering status in the realm of electronics.

During his time at Centralab, Jack Kilby's career was distinguished by significant achievements, including obtaining several patents¹¹¹. One patent was for a miniaturized solid state hearing aid, illustrating Kilby's early exploration into the realm of compact electronic design. Reflecting on his tenure at Centralab, Kilby often spoke of how this period ignited his passion for invention¹¹².

The new job at TI was ideal. TI was one of the first companies to commercialize transistor-based devices, and they hired Kilby into a special unit that was working on miniaturizing solid state electronic devices¹¹³. He had a very strong theoretical background, liked building things with his own hands, and was infused with the spirit of invention.

Jack's Epiphany

In the humid Dallas July of 1958 Jack Kilby was 34 years old and two months into the new job at TI. His boss and colleagues were away on vacation, so he worked alone in the unairconditioned lab on the problem that was holding back the miniaturization of electronic devices¹¹⁴.

Jack Kilby's demonstration of the first working integrated circuit (IC) in 1958 revolutionized the field of microelectronics¹¹⁵. Instead of using individual transistors, resistors, and capacitors to form a circuit, Kilby's first IC combined a transistor, a capacitor, and the equivalent of three resistors all on the same chip¹¹⁶. This first IC was a simple form of a digital memory for one binary digit of information.

Jack solved the interconnection problem that was holding back miniaturization, and the embodiment of his solution is now inches from you and nearby to almost everyone you know. In fact, the man-made result of this idea is more numerous and omnipresent than blades of grass on the planet¹¹⁷. This product of Jack's idea orbits Earth, is on the moon and Mars and has even

¹¹⁰ Frye, John T. *Printed Circuits*. p. 46

¹¹¹ Kilby, Jack S. *Modular Electrical Unit*, US Patent No. 3,052,822

¹¹² Reid, Thomas p. 72

¹¹³ Pirtle III, Caleb. *Engineering the world*. p. 81

¹¹⁴ Reid, Thomas *The chip* p.6

¹¹⁵ Reid, Thomas p.73

¹¹⁶ Jack Kilby's Integrated Circuit. National Museum of American History. Catalog Number 1987.0487.320

¹¹⁷ Method of estimate: There are approximately 4.6 quintillion blades of grass on the Earth, by estimating 100 blades of grass per square inch and that 30% of the Earth's land area is covered by grass. There are over 50 quintillion transistors with over 10 billion integrated circuits and an average of over 5 billion

traveled 22 billion miles away—past the edge of the solar system on the Voyager spacecraft¹¹⁸. This invention of the microchip triggered an exponentially accelerating cascade that brought about the modern world¹¹⁹.

Building a circuit is like constructing a sentence. There are standard components with a particular arrangement—in a sentence there are nouns, verbs, adjectives—in a circuit there are transistors, resistors, capacitors and diodes. In order to construct the circuit on a single substrate, Kilby had to invent a method to construct each of the required component types from the same material. In his case this substrate was the semiconducting element from the periodic table called germanium. The concept of integrating multiple components on a single substrate fundamentally shifted the trajectory of electronics development. Without this leap, electronic advancements would have been constrained by the limitations of connecting individual components.

Competition

About six months after Kilby, Robert Noyce, working at Fairchild Semiconductor, developed a similar technology¹²⁰. The story of Robert Noyce, who is often referred to as the "Mayor of Silicon Valley"¹²¹ is deeply entwined with Kilby's, illustrating the complex nature of technological progress. Amazingly, Kilby and Noyce had solved separate parts of the problem.

While Kilby's novel idea was that all of the circuit components could be made from the same material and built on the same substrate, Noyce had invented a new way to connect the components together. With the method invented by Noyce, the wires between the components could be "printed" onto the substrate in a series of layers called photolithography. This is analogous to the way modern color laser printing works, where each color is added to the page sequentially, resulting in the desired mixture of colors at each spot on the page. In photolithography there are a series of masks for each layer, and the materials in the layers are metals and glass (insulators).

Also, Kilby used the element germanium as a substrate, and Noyce used the element silicon as the semiconductor substrate. Silicon is one of the most common elements on the Earth. For example, desert sand is about 50% silicon, and glass is a combination of silicon and oxygen (silicon dioxide). Germanium, in contrast, is relatively rare¹²². In short, the process proposed by Noyce allowed for more reliable and easier mass production of microchips¹²³.

Robert Noyce's impact transcends his technical contributions—he was instrumental in fostering the collaborative, innovative culture of Silicon Valley. As a co-founder of Fairchild Semiconductor and later Intel, Noyce helped create an environment that encouraged

transistors per device. Therefore there are about 10 times more transistors in integrated circuits than blades of grass.

¹¹⁸ Ramesh, Shreya *Voyager: Venturing Outside the Solar System*. Pp. 13-6

¹¹⁹ Zygmunt, Jeffrey *Microchip* p. 170

¹²⁰ Reid, Thomas *The chip* p. 93

¹²¹ Berlin, Leslie. *The man behind the microchip* p. 87

¹²² Höll, R. et. al.. *Metallogenesis of germanium—A review*. p.145

¹²³ Reid, Thomas *The chip* p. 94

entrepreneurial spirit and risk-taking¹²⁴, which became the hallmark of Silicon Valley and a driving force behind the tech industry's explosive growth.

The simultaneous invention led to a 'patent war' between Texas Instruments and Fairchild Semiconductor. Eventually, the dispute was settled with a cross-licensing agreement, acknowledging the contributions of both Kilby and Noyce. This settlement was a landmark in the industry, setting a precedent for handling intellectual property disputes in rapidly evolving technological fields¹²⁵.

The stories of Kilby and Noyce exemplify how technological breakthroughs often result from the interaction of multiple minds. While Kilby laid the groundwork with his initial invention, Noyce's refinements and business acumen played an equally crucial role in bringing the integrated circuit to the global market, revolutionizing technology as we know it.

While it is likely that some form of integrated circuitry would have been developed eventually, given the technological trends and needs of the time, Kilby's invention was a definitive and pivotal moment in history. It broke the existing paradigm, which was to further improve individual components, and set a new course for the future of technology, making it more than just a step in a continuous progress.

Moore's Law

Kilby's invention, and the subsequent development of a similar technology by Robert Noyce at Fairchild Semiconductor, marked the inception of the microchip era. These early ICs were primitive by today's standards, containing only a few transistors. However, they represented a monumental leap from the bulky, vacuum-tube-based electronics that dominated the time.

Since Kilby's invention, the number of transistors that could be placed on a single chip doubled approximately every two years, a trend famously encapsulated by Moore's Law¹²⁶. This law was postulated in 1965 by Gordon Moore, who co-founded Intel with Robert Noyce. Now, the 21st century has seen the integration of microchips into almost every aspect of daily life. The rise of smartphones, smart homes, and the Internet of Things (IoT) has created an interconnected world where microchips play a vital role in communication, automation, and artificial intelligence¹²⁷. Today's chips, featuring billions of transistors, are not just components in computers; they are the bedrock of modern technology, powering everything from cars to medical devices to smart city infrastructures¹²⁸.

A modern computer chip like the Apple M3 Max that is manufactured with the new three nanometer technology, has 92 billion transistors, is less than four cubic inches in size, and costs less than \$500¹²⁹. The modern processor has two million times more transistors, is almost a million times smaller, and is over 100,000 times less expensive. This miniaturization is possible

¹²⁴ Berlin, Leslie. *The man behind the microchip* p. 87

¹²⁵ Reid, Thomas p. 117

¹²⁶ Mack, C. A.. Fifty years of Moore's law. p. 202

¹²⁷ Miller, Chris. *Chip war* p. 238

¹²⁸ Miller, Chris. pp. 279-80

¹²⁹ Apple Computer Corp. Press Release. Apple unveils M3

because all of these transistors and components are “printed” onto a single substrate. The combined transistors, wires, and other components on a single substrate is called an IC, and this was Jack Kilby’s invention.

The development of the three nanometer manufacturing process in 2022 by Samsung and TSMC, which was used to fabricate the Apple M3 processors, marks a significant milestone in microchip technology¹³⁰. This cutting-edge technique dramatically reduces transistor size, enabling the placement of billions more transistors on a chip. The result is not just a boost in processing power but also enhanced energy efficiency, a crucial aspect in the age of environmental consciousness and sustainable technology¹³¹.

Kilby’s role in bringing about this microchip revolution was celebrated with many awards, including the Nobel Prize in Physics, awarded to Jack Kilby in 2000¹³². In his Nobel lecture, Kilby highlighted the importance of the work of Robert Noyce, who did not share the prize as he had passed away a decade earlier.

Kilby continued to invent throughout his career. He has the initial patent for the electronic, hand-held, solid-state calculator, and fifty-two additional patents¹³³. In addition to the Nobel prize, he received the National Medal of Science and his picture hangs in the Inventors Hall of Fame in Washington D. C. between those of Henry Ford and Ernest Lawrence, inventor of the cyclotron¹³⁴.

Currently, the relentless miniaturization of microchip components is facing fundamental physical constraints¹³⁵. As transistors approach atomic sizes, quantum effects and other challenges emerge, prompting a shift in focus from mere size reduction to enhancing overall chip performance and efficiency¹³⁶. On one hand it appears improving performance of electronics has reached a roadblock, but Kilby and Noyce provide inspiration, and their breakthrough demonstrates that solutions are never too far out of reach.

The race forward

There is a global competition to lead in the advancement of microchip technology, often referred to as the “chip wars,” and it is more than a race for technological supremacy; it is a strategic battle with significant economic, political, and security implications. Central to this contest is the role of microchips in advancing artificial intelligence (AI), and legislative actions like the CHIPS Act (Creating Helpful Incentives to Produce Semiconductors)¹³⁷, which demonstrate concerted efforts by nations to gain or maintain a leading edge in this critical field.

The “chip wars” are a reflection of the pivotal role that semiconductor technology and Kilby’s invention plays in modern economies and national security. As the backbone of

¹³⁰ Miller, Chris. *Chip war* p. 331

¹³¹ Recce, Alexander. *Computing a lower carbon footprint*. pp. 9-10

¹³² Kilby, Jack St Clair. *Turning potential into realities*. Nobel Lecture pp. 482-89

¹³³ Eckstein, Peter. *Jack Kilby (1923-2005)* p. 91

¹³⁴ Reid, Thomas *The chip* p. 62

¹³⁵ Cao, W. et. al. *The future transistors* p. 501

¹³⁶ Cao, W. et. al. p. 503

¹³⁷ H.R. 4346 Law enacted by 117th Congress in 2022 authorizes \$280 billion in stimulus

everything from consumer electronics to advanced weapons systems, control over microchip production and innovation is a significant strategic advantage. The race involves technological powerhouses like the United States, China, South Korea, Taiwan, and European Union nations, each striving to secure a self-reliant and dominant position in semiconductor manufacturing and design¹³⁸.

Conclusion

Microchips are indispensable in the field of AI. The processing power and efficiency of advanced chips determine the speed and capability of AI algorithms. As AI applications become increasingly sophisticated, the demand for high-performance, energy-efficient microchips continues to soar. This demand has spurred innovations in chip design specifically tailored for AI tasks, including application-specific neural network chips and deep learning processors.

The future of the chip wars is likely to be characterized by rapid innovation, shifting alliances, and significant investments in research and development. As nations recognize the strategic value of microchip technology, the race is set to intensify, with implications for global technological leadership and economic power.

The advances in integrated circuits, as experienced by the use of technology, have occurred for so many decades that for most people there is an assumed undercurrent of rapid progress for their entire life. There are significant engineering challenges in continuing this progress, and growing geopolitical efforts to lead this innovation, so the future is not clear. However the human inspiration that drove Jack Kilby to invent on those hot days in the Texas lab in 1958 gives us faith that another engineer (or two) will make it work. As we head into a future where rival nations and companies strive to extend the frontiers of computing, we can expect to find positive competition and collaboration between engineers at the source, as discussed in this analysis of the birth of the microchip.

¹³⁸ Miller, Chris. *Chip war* p. 324

Works Cited

- Apple Computer Corp. Press Release. Apple unveils M3, M3 Pro, and M3 Max, the most advanced chips for a personal computer. October 30, 2023.
- Bashe, Charles J., Lyle R. Johnson, John H. Palmer, and Emerson W. Pugh, eds. *IBM's early computers*. MIT press, 1986.
- Berlin, Leslie. *The man behind the microchip: Robert Noyce and the invention of Silicon Valley*. Oxford University Press, 2005.
- Cao, W., Bu, H., Vinet, M., Cao, M., Takagi, S., Hwang, S., ... & Banerjee, K. "The future transistors." *Nature* 620, no. 7974 (2023): 501-515.
- Eckstein, Peter. "Jack Kilby (1923-2005)." *IEEE Annals of the History of Computing* 29, no. 01 (2007): 90-95.
- Frye, John T. "Printed Circuits". *Radio News*, no. 12 (1948), 46
- Halberstam, David. *The fifties*. Ballantine Books, 1994.
- Höll, R., M. Kling, and E. Schroll. "Metallogenesis of germanium—A review." *Ore Geology Reviews* 30, no. 3-4 (2007): 145-180.
- Jack Kilby's Integrated Circuit. National Museum of American History. Catalog Number 1987.0487.320. Washington D.C. Credit Line: From Texas Instruments
- Kilby, Jack S. Modular Electrical Unit, US Patent No. 3,052,822. Issued September 4, 1962.
- Kilby, Jack S. Miniaturized electronic circuits, US Patent No. 3,138,743. Issued June 23, 1964.
- Kilby, Jack S. "Turning potential into realities: The invention of the integrated circuit (Nobel lecture)." *ChemPhysChem* 2, no. 8-9 (2001): 482-489.
- Mack, C. A. Fifty years of Moore's law. *IEEE Transactions on semiconductor manufacturing*, 24(2), (2011): 202-207.
- Miller, Chris. *Chip war: the fight for the world's most critical technology*. Simon and Schuster, 2022.
- Pirtle III, Caleb. "Engineering the world: Stories from the first 75 years of Texas Instruments." (No Title) (2005)
- Ramesh, Shreya. "Voyager: Venturing Outside the Solar System." *Berkeley Scientific Journal* 27, no. 2 (2023): 13-16.
- Recce, Alexander M. Computing a lower carbon footprint. Pingry Community Research (PCR). Fall 2023, 9-10.
- Reid, Thomas R. *The chip: How two Americans invented the microchip and launched a revolution*. Random House Trade Paperbacks, 2001.
- Zygmunt, Jeffrey. *Microchip: An idea, its genesis, and the revolution it created*. Perseus Publishing, 20

Bridging Scientific Disciplines & Cultures: A Multidisciplinary and Multicultural Approach to Palliative Care By Hannah Yin

Abstract

- Brief overview of the importance of palliative care
- Summary of key scientific and technical elements
- Outline of the research paper's goals and structure

Palliative care, also known as end-of-life care, is a very important field in healthcare, focusing on improving the quality of life for patients with serious illnesses by addressing their physical, psychological, social, and spiritual needs. This research article goes through the multidisciplinary and multicultural factors of palliative care, making sure to highlight the importance of the mix of different scientific and technical fields needed to provide thorough care. Some of these key areas include medical and nursing science, pharmacological approaches, psychological and emotional support, nutritional and social support, ethical considerations, technological advancements, and physical and occupational therapy. This article aims to explore how all of these elements come together to improve palliative care, with an emphasis on the influence of cultural diversity on end-of-life care. This paper includes an in-depth analysis of each field's contribution to palliative care and the multicultural approach in improving patient outcomes and satisfaction.

Introduction

- Definition and goals of palliative care
- Importance of a multidisciplinary approach
- Importance of a multicultural approach
- Overview of the patient population and typical scenarios in palliative care

Palliative care is a specialized medical approach that's designed for people with complex, ongoing illnesses, with its main goal being to relieve the pain, symptoms, and stress associated with serious medical conditions, so it enhances the quality of life for both patients and their families. People who are in palliative care are dealing with serious illnesses such as cancer, heart failure, chronic obstructive pulmonary disease (COPD), kidney disease, Alzheimer's, Parkinson's, Amyotrophic Lateral Sclerosis (ALS), and many more. Palliative care makes sure to address the physical, psychological, social, and spiritual needs of patients, so that healthcare workers can provide holistic support through all stages of illness.

However, a multidisciplinary approach and multicultural approach is fundamental to make sure the care given is as effective as possible. This approach involves a unified team of healthcare professionals, including doctors, nurses, social workers, psychologists, and spiritual advisors, who come together to address the diverse needs of patients. By integrating various expertise, the palliative care team makes sure comprehensive and personalized care plans are created that improve patient outcomes and their quality of life. In our increasingly diverse society, a multicultural approach to palliative care is also very important because cultural beliefs and practices significantly influence how patients in care perceive their illness, make medical decisions, and respond to their treatment. Being able to understand and respect these cultural differences is crucial for providing effective and compassionate care to patients. This multicultural approach involves being sensitive to these differences, ensuring that care plans are tailored to meet the unique needs of patients from various cultural backgrounds because this not only enhances patient satisfaction and trust but also improves the overall efficiency of the care that is given.

The patient population in palliative care is diverse, including individuals with a wide range of complex illnesses such as cancer, heart failure, chronic obstructive pulmonary disease (COPD), kidney disease, Alzheimer's, Parkinson's, and Amyotrophic Lateral Sclerosis (ALS). Patients often face a variety of symptoms, including pain, depression, shortness of breath, fatigue, constipation, nausea, loss of appetite, difficulty sleeping, and anxiety. Addressing these symptoms requires a personalized approach for each patient, taking into account their cultural background, preferences, and goals. Studies have also shown that patients receiving palliative care not only experience improved quality of life but also tend to live longer than those who do not receive this care.

In conclusion, the mix of multidisciplinary and multicultural approaches in end-of-life care is essential for providing high-quality, compassionate care to diverse patient populations. This research article will explore the contributions of various scientific disciplines to palliative care and the importance of cultural sensitivity in improving patient outcomes and satisfaction before they pass away.

1. Medical and Nursing Science in Palliative Care

1.1 Medical Knowledge

- Pathophysiology of common illnesses in palliative care
- Basic principles of symptom management
-

In palliative care, an understanding of the pathophysiology underlying common illnesses is important for helping one provide effective care and improve patient outcomes. [1] One common illness that many in palliative care have is cancer. Cancer is characterized by uncontrolled cell growth and proliferation, which often leads to the formation of tumors that can

invade nearby tissues and spread to distant sites in the body. The pathophysiology of cancer includes genetic mutations that disrupt normal cell cycle regulation and apoptosis, which results in sustained proliferation and the evasion of the immune system's surveillance. In palliative care, addressing cancer-related symptoms such as pain, fatigue, and cachexia requires targeted therapies that reduce the tumor burden and enhance patient comfort.

Another common complex illness that people in palliative care typically have is heart failure. Heart failure occurs when the heart is unable to pump blood efficiently to meet the body's demands, which leads to fluid retention, organ congestion, and reduced tissue perfusion. Pathophysiologically, heart failure can result from myocardial damage, chronic hypertension, or valvular disorders, which impairs your cardiac function over time. Palliative care for heart failure focuses on managing symptoms such as dyspnea, edema, and fatigue through medications, lifestyle modifications, and supportive therapies that are aimed to optimize the patient's quality of life.

Chronic obstructive pulmonary disorder (COPD) includes chronic bronchitis and emphysema, and is characterized by progressive airflow limitation and respiratory symptoms in the human body. Some pathophysiological changes also include airway inflammation, mucus hypersecretion, and structural changes in lung tissue, which leads to airflow obstruction and abnormal gas exchanges. Palliative care interventions for COPD emphasize symptom relief through bronchodilators, oxygen therapy, pulmonary rehabilitation, and end-of-life planning to support patients and families that are facing advanced disease stages.

Neurodegenerative diseases like Alzheimer's, Parkinson's, and amyotrophic lateral sclerosis (ALS) involve progressive degeneration of neurons and neuronal networks, which leads to cognitive decline, motor dysfunction, and the eventual loss of independence for patients. Pathophysiological mechanisms include protein aggregation, oxidative stress, and neuroinflammation, which contribute to neuronal damage and functional impairment. Palliative care for neurodegenerative diseases focuses on managing symptoms such as pain, dysphagia, spasticity, and behavioral changes while supporting caregivers in navigating the complex care needs for these illnesses.

Effective symptom management in palliative care is guided by the foundational principles that prioritize patient comfort, holistic assessment, and personalized interventions tailored to the individual needs of each patient. Successful symptom management begins with a thorough assessment encompassing physical, psychological, social, and spiritual dimensions of patient experience. Tailoring care plans to each patient's unique needs and preferences is fundamental in palliative care, because individualized plans use evidence-based practices with patient-centered goals, which creates collaboration between healthcare providers, patients, and families to optimize symptom relief and promote shared decision-making throughout the care plan.

1.2 Nursing Science

- Role of nursing in direct patient care

- Development and implementation of personalized care plans

Nursing science is crucial to the act of direct patient care, particularly in the area of palliative care. Nurses are at the lead of patient interactions with palliative care, and they offer not only medical assistance but also emotional and social support to patients and families. Their responsibilities include a diverse range of tasks that are aimed at improving the quality of life for patients who have serious illnesses and are at end-of-life. Palliative care nurses are trained to monitor and manage pain, lessen physical symptoms, and deliver medications effectively. However, their role goes far beyond physical care; they provide emotional support to patients and their families, making sure that they are well-informed about the patient's condition and care options. This holistic approach is very important in palliative care because the focus of the care is to enhance the quality of life for patients that are facing life-limiting conditions, and are aware that their time is coming to an end. Nurses are also responsible for educating and advising families, helping them go through the complexities of the healthcare system, and they offer practical advice on managing the patient's condition at home. This support is important in helping families cope with the stress and emotional trauma associated with having a family member in palliative care.

One of the main skills of nursing in palliative care is the development and implementation of personalized care plans. These care plans are personalized to meet the unique needs of each patient, and they take into consideration their medical condition, personal preferences, cultural background, family dynamics, religion, etc. These personalized care plans are created through extensive assessments that include evaluating the patient's physical, emotional, social, and spiritual needs. Nurses then collaborate with an interdisciplinary team, that typically includes doctors, social workers, chaplains, and other healthcare professionals, to create a care plan that addresses all the aspects of the patient's well-being to increase their quality of life. The implementation of these care plans needs careful coordination and communication, and nurses play an important role in making sure that all the members of the care team are informed so that the care plan is consistently followed. Nurses also monitor the patient's progress and make adjustments to the care plan as needed, so that it remains relevant to the patient's needs and continues to be effective throughout the course of the illness. Additionally, nurses advocate for their patients, making sure that their wishes and needs are respected and met which is really important in palliative care because patients typically have specific wishes regarding their end-of-life care, and by supporting this patient-centered care, nurses help patients receive the care that is aligned with their values and goals.

To summarize, nursing science in palliative care involves a diverse approach to direct patient care and the development and implementation of personalized care plans. Nurses' expertise in managing symptoms, providing emotional support, educating families, and advocating for patients is important in delivering high-quality palliative care that helps enhance the quality of life for all patients.

2. Pharmacological Approaches

2.1 Pain Management

- Types of pain typical in palliative patients
- Overview of opioids and non-opioid pain management strategies

Pain in palliative care is a subjective experience, and you commonly encounter this in patients with advanced diseases such as metastatic cancer, multiple sclerosis, cardiac disease, diabetes mellitus, and in chronic illnesses like osteoarthritis. Pain can be broadly classified into two categories: acute and chronic pain. Acute pain is usually the result of a specific injury, such as a fracture or bruise, and the pain usually goes away as the injury heals. This type of pain is often managed by generalist providers unless complicated by other factors, such as a recent diagnosis of a life-limiting illness or additional complex symptoms. Chronic pain, on the other hand, goes on for more than three months after the initial injury has healed, and can be increased by behavioral components like depression and anxiety, which develop as the individual adapts to their pain experience. For patients with nonmalignant chronic pain, a pain management specialist may be necessary, especially if there are concerns about potential substance abuse, however for those with chronic pain related to serious illnesses, such as malignancy, a palliative care specialist is often needed.

Pain pathophysiology can be divided into two distinct types: nociceptive and neuropathic pain. Nociceptive pain comes from tissue injury, while neuropathic pain results from nerve injury. There are various non-pharmacologic methods to manage pain, which can be used alone or alongside medication depending on the severity and type of pain. Techniques include repositioning, hot and/or cold packs, music therapy, physical or occupational therapy, relaxation techniques, transcutaneous electrical nerve stimulation (TENS), and acupuncture. These methods can provide lots of relief and are essential components of a comprehensive pain management plan for those in palliative care. In 1986, the World Health Organization (WHO) introduced a stepped approach for managing cancer-related pain, known as the WHO pain ladder. This approach has since been adapted to manage all types of pain, including nonmalignant pain such as osteoarthritis and chronic back pain. The ladder suggests starting with non-opioid analgesics, such as acetaminophen and NSAIDs, which have a ceiling effect, meaning increasing doses beyond a certain point does not enhance pain relief but raises the risk of side effects. For instance, doses of acetaminophen exceeding 4 grams per day can lead to liver damage, and NSAIDs, beyond a certain dose, increase the risk of gastrointestinal, renal, and cardiovascular side effects.

Opioid analgesics do not have a ceiling effect, which allows doses to be increased to achieve pain relief without a proportional increase in side effects. However, it comes with its own set of challenges. Side effects such as sedation, decreased respiratory drive, nausea, vomiting, pruritus, and constipation are common and must be managed proactively. For instance,

opioid-induced constipation (OIC) requires a concurrent bowel regimen to be initiated with opioid therapy. Sedation, occurring in 20–60% of patients when starting or increasing opioid doses, usually decreases after a few days, however, persistent or severe sedation may need dose adjustments or specialist referral. [2] Decreased respiratory drive is life-threatening and requires immediate intervention, especially in opioid-naïve individuals. Nausea and vomiting, affecting 10–40% of patients, typically resolve within a week, while pruritus, seen in 2–10% of patients, can be managed with antihistamines or opioid rotation. Constipation is a persistent issue and requires ongoing management with laxatives and stool softeners. Effective pain management requires careful consideration of dosing and titration. Oral administration is preferred, and opioid-naïve patients should start with short-acting opioids to assess their response/reaction to the medicine. Pain occurring around the clock necessitates continuous dosing, and anticipating side effects, especially constipation, is crucial, which is why nonpharmacologic approaches should complement pharmacologic management to potentially reduce medication requirements.

Understanding the difference between tolerance, physical dependence, pseudoaddiction, and addiction is vital for appropriate pain management. Tolerance involves needing higher doses for the same analgesic effect, while physical dependence results in withdrawal symptoms if the medication is abruptly discontinued. Pseudoaddiction mimics addiction behaviors driven by unrelieved pain and resolves once pain is adequately managed. True addiction involves maladaptive behaviors, including impaired control over drug use, compulsive use, and continued use despite harm.

2.2 Managing Other Common Symptoms

- Antiemetics for nausea
- Drugs for anxiety and sleep disturbances

Nausea is a widespread and distressing symptom in healthcare that can appear from various causes, ranging from simple reactions to ingested toxins to complex signals activated by motion, medications, anesthesia, stress, pregnancy, psychiatric disorders, or fear. Multiple neurohumoral pathways may be involved in its appearance, and managing nausea requires careful evaluation by the healthcare provider, as the therapy for chronic nausea, like one resulting from chemotherapy, can differ significantly from the treatment of acute postoperative nausea and vomiting. The mechanism of action of antiemetics involves several brain regions, including the area postrema in the floor of the fourth ventricle, the nucleus tractus solitarius in the medulla, and various motor nuclei. These regions interact with vagal afferents and efferents to the cortex, hypothalamus, and limbic regions, influencing how nausea is perceived. Multiple chemoreceptors, including muscarinic M1, dopaminergic D2, serotonin 5-HT₃, neurokinin-1 (NK-1), and histamine H₁, are involved in these pathways. As a result, most antiemetics can be categorized based on their mechanism of action.

Serotonin-receptor antagonists, such as ondansetron and granisetron, work by blocking serotonin from interacting with the 5-HT₃ receptor. These medications are commonly used and are available in both intravenous (IV) and oral (PO) preparations, with side effects including headache, dizziness, and constipation. Glucocorticoids, like dexamethasone, are also used, particularly in chemotherapy and postoperative settings, although their mechanism of action is not entirely understood. Anticholinergics, such as scopolamine, are mainly used to treat motion sickness and are administered transdermally. NK receptor antagonists, like aprepitant, prevent the release of substance P, a vomiting inducer, and are used in both PO and IV forms.

Dopamine receptor antagonists, such as prochlorperazine and chlorpromazine, block the D₂ receptor and are used in various formulations, including IV, PO, and rectal (PR). These medications can cause side effects like dizziness and extrapyramidal symptoms. Butyrophenones, such as droperidol and haloperidol, also antagonize the D₂ receptor but have fallen out of favor due to their side effect profiles. Benzamides, like metoclopramide, antagonize the D₂ receptor at low doses and the 5-HT₃ receptor at higher doses, serving as promotility agents to reduce nausea and vomiting. [3] Cannabinoids, such as nabilone and dronabinol, have shown some benefit but come with significant side effects like vertigo and dysphoria.

Antihistamines, such as diphenhydramine and promethazine, are widely available and generally well-tolerated, with sedation being a common side effect. Monitoring and contraindications for these medications vary based on the class and mechanism of action. For instance, serotonin-receptor antagonists should be used cautiously in patients with prolonged QTc, while glucocorticoids are contraindicated in those with systemic fungal infections. Proper monitoring and careful selection of antiemetics can enhance patient outcomes, especially when coordinated by an interprofessional healthcare team.

Anxiety and sleep disturbances are common symptoms in palliative care patients, significantly impacting their quality of life. Effective management requires a comprehensive approach, often involving both pharmacologic and non-pharmacologic strategies. Pharmacologic treatments for anxiety include benzodiazepines, such as lorazepam and diazepam, which enhance the effect of the neurotransmitter GABA at the GABA-A receptor, providing a calming effect. These medications are effective but can lead to dependency and should be used cautiously. For sleep disturbances, medications such as zolpidem, a non-benzodiazepine hypnotic, are often prescribed. Zolpidem works by enhancing GABA activity, similar to benzodiazepines, but with a different chemical structure. It is effective for short-term treatment of insomnia but can cause side effects like dizziness and daytime drowsiness. Another option is trazodone, an antidepressant with sedative properties, which can be particularly useful for patients who experience both depression and insomnia. Trazodone's primary mechanism is the inhibition of serotonin reuptake, along with blocking histamine and alpha-1 adrenergic receptors, contributing to its sedative effects.

Non-pharmacologic strategies are also crucial in managing anxiety and sleep disturbances. Cognitive-behavioral therapy (CBT) has been shown to be effective in reducing anxiety and improving sleep quality. Relaxation techniques, such as progressive muscle

relaxation and mindfulness meditation, can help alleviate anxiety and promote better sleep, and creating a conducive sleep environment and establishing a regular sleep routine are also important non-pharmacologic interventions. In summary, managing common symptoms such as nausea, anxiety, and sleep disturbances in palliative care involves a combination of pharmacologic and non-pharmacologic approaches. An interprofessional team, including physicians, nurses, pharmacists, and mental health professionals, can provide comprehensive care, ensuring that patients receive the most effective and individualized treatment for their symptoms.

3. Psychological and Emotional Support

3.1 Psychology and Psychiatry

- Managing psychological symptoms: depression, anxiety
- Techniques and therapies for emotional support

People often assume that those diagnosed with a life-threatening illness will automatically experience depression. Though it is true because patients facing serious illnesses are more likely to suffer from depression or anxiety compared to healthy individuals. For example, one study of terminally ill cancer patients found that at least 17% were clinically depressed, while other studies show even higher percentages of people with terminal illnesses meeting the criteria for major depression. It is essential to distinguish between natural emotional responses to severe illness and true clinical depression, which goes beyond typical sadness and anxiety. Experts point out that it's perfectly natural to feel scared, sad, and anxious about death and the dying process, and these feelings arise because these individuals experience something they have never faced before. The palliative care team plays a crucial role in helping patients work through these emotions, however, true clinical depression differs from normal grieving and anxiety processes, and it's often underdiagnosed but should be identified and treated.

Signs of clinical depression in patients with life-threatening illnesses include a lack of interest in previously enjoyed activities, significant changes in sleeping or eating habits, withdrawal from friends and family, and serious thoughts or discussions about suicide. Recognizing these symptoms is important for both patients and caregivers because if these signs are present, it is important to seek help from a doctor or care team member, and immediate medical assistance should be sought if there are thoughts of suicide. Clinical depression in palliative care patients can be treated effectively with things like antidepressant treatments that are as effective in palliative care patients as in the general population. However the most effective treatments usually combine short-term psychotherapy with antidepressant medications. For patients experiencing grief and anxiety that do not meet the criteria for full-scale depression, medications might not be necessary, but emotional support remains crucial. Palliative care professionals emphasize the importance of educating both the patient and their caregivers about

what to expect, and much of the anxiety and negative emotions associated with a life-threatening illness come from feelings of helplessness and uncertainty. By answering questions and providing talk therapy, social workers and other members of the palliative care team help patients live as well as possible for as long as possible.

Communication plays a major role in reducing anxiety near the end of life. Both the dying person and their loved ones often avoid discussing what is happening due to fears about how these conversations will affect each other, but the palliative care team can facilitate open, calm discussions, decreasing much of the anxiety and making the process easier for everyone involved. Effective management of psychological symptoms in palliative care involves a range of techniques and therapies tailored to patients' individual needs. Cognitive-behavioral therapy (CBT) [4] is widely used to address depression and anxiety by helping patients reframe negative thoughts and develop coping strategies, and supportive psychotherapy provides a space for patients to express their fears and concerns, creating emotional resilience. Mindfulness and relaxation techniques, such as meditation and deep-breathing exercises, can also reduce anxiety and improve overall well-being because these practices help patients focus on the present moment, reducing the preoccupation with future uncertainties and reducing stress.

Family therapy and counseling can be beneficial, as they involve the patient's support system in the therapeutic process, addressing relational dynamics and enhancing communication. This holistic approach ensures that both patients and their loved ones receive the emotional support they need. Pharmacological interventions, including antidepressants and anxiolytics, may be prescribed for patients with severe symptoms, and these medications can be particularly useful when psychological therapies alone are insufficient. However regular monitoring and adjustment of these medications are necessary to manage side effects and ensure their effectiveness. In summary, managing psychological symptoms in palliative care requires a multifaceted approach. By combining pharmacological treatments, psychological therapies, and supportive care, palliative care teams can effectively address depression and anxiety, helping patients navigate the emotional challenges of life-threatening illnesses.

3.2 Role of Counseling, Multicultural & Spiritual Care

- Importance of addressing existential questions and spiritual needs
- Case studies on effective psychological interventions

In palliative care, addressing the existential and spiritual needs of patients is important for providing holistic care that respects the patient's dignity, beliefs, and cultural background. The biopsychosocial-spiritual model shows the importance of viewing the patient as a whole person, where spiritual and existential well-being are as major as physical and psychological health. Spirituality, different from religion, usually serves as a source of comfort, meaning, and connection for patients facing the end of life, and neglecting this factor can result in increased psychological and physical distress.

Spiritual distress, often characterized by feelings of alienation, loss of meaning, and inner turmoil, can increase other physical and psychological symptoms such as pain, anxiety, and depression. Because of this, addressing spiritual needs is not just an ethical obligation, but a clinical necessity to improve the patient's quality of life. Healthcare professionals, including psychologists, should be equipped to engage in conversations about spirituality, understanding its nuances and how it intersects with the patient's cultural and religious background. This approach aligns with the need for culturally sensitive care, where the diverse beliefs and values of each patient are acknowledged and respected.

One notable case involved a middle-aged woman with advanced cancer who experienced profound spiritual distress, [5] being expressed as a loss of meaning and severe depression. Her healthcare team, recognizing the inadequacy of only physical interventions, involved a psychologist trained in palliative care. Through a series of narrative therapy sessions, the psychologist facilitated the patient's exploration of her life story, focusing on meaningful relationships and past experiences that had given her life purpose. The patient was also encouraged to express her spiritual concerns, which were addressed in collaboration with a chaplain. This multidisciplinary approach significantly reduced her spiritual distress, improved her mood, and helped her find peace in her remaining time. Another case involved an elderly man in a hospice setting who struggled with feelings of guilt and fear related to his religious beliefs about the afterlife. He expressed a deep sense of anxiety about dying, fearing divine punishment. The psychologist utilized an explicit integration model, where religious themes were openly discussed in therapy. The patient was guided through a process of self-forgiveness and reconciliation with his beliefs, and a spiritual care worker provided religious rituals that aligned with his faith. This intervention alleviated his anxiety and allowed him to approach his end-of-life experience with a sense of serenity and acceptance.

These cases highlight the effectiveness of integrating spiritual care into psychological interventions in palliative settings. By addressing the spiritual and existential dimensions of suffering, psychologists can play a life-changing role in enhancing the overall well-being of patients. Furthermore, these interventions reveal the importance of a culturally sensitive approach, where the patient's spiritual and religious beliefs are respected and incorporated into the care plan. This not only honors the patient's autonomy but also fosters a deeper sense of peace and fulfillment as they approach the end of life.

4. Nutritional Science and Social Support

4.1 Nutritional Management

- Common nutritional challenges and interventions
- Impact of nutritional support on quality of life

In palliative care, reduced oral intake or malnutrition is a frequent and significant challenge that requires careful management. As diseases progress, symptoms such as fatigue,

muscular weakness, and dysphagia often worsen, making eating and drinking increasingly difficult for patients. This not only impacts them physically but also has major psychological, social, and cultural implications. Food is more than just sustenance; it is closely associated with health, well-being, and social interaction. [6] For patients and their caregivers, the inability to eat can create fear and concern, deeply affecting the patient's quality of life, including their relationships with loved ones. The nutritional objective in palliative care should, therefore, focus on quality of life, emphasizing comfort, symptom relief, and the enjoyment of food. While it is crucial to prevent or treat malnutrition, nutritional support measures must be balanced against their potential to impair quality of life. The Care Quality Commission mandates that healthcare establishments meet their patients' nutritional needs, requiring thorough nutritional assessments. However, the complex and varied needs of palliative patients call for person-centered, individualized nutrition plans that can adapt to changes in the patient's condition.

Effective nutritional interventions may include modifying food or drink to address specific issues like dysphagia, as well as using oral nutritional supplements. In the later stages of terminal illness, the goal of nutrition therapy typically shifts from maintaining energy intake or reversing malnutrition to focusing on comfort and quality of life. Aggressive feeding may not be appropriate, especially if it causes discomfort or anxiety. However, in some cases, even small or symbolic interactions with food and drink can be meaningful for patients, helping them feel connected to others. There may come a point where maintaining adequate nutrition is only possible through artificial feeding, such as medically assisted nutrition and hydration. However, the decision to use artificial feeding should be based on the perceived benefits, harms, risks, and burdens for each individual patient. A Cochrane review highlighted the lack of randomized controlled trials to support specific recommendations for artificial hydration and nutrition in palliative care, revealing the need for careful, individualized decision-making.

Ultimately, the primary goal in terminal care is the comfort of the patient. Decisions about whether to use artificial feeding should prioritize this goal. The General Medical Council (GMC) provides guidance on meeting nutrition and hydration needs in end-of-life care, but emphasizes that oral nutrition and hydration should always be supported whenever possible. It is important to offer food and drink by mouth as part of basic care, and any decision to withhold clinically-assisted nutrition or hydration should be based on the lack of beneficial effect and not on the preferences of clinical staff. Nutritional support in palliative care has a significant impact on the quality of life of patients. As symptoms progress and the ability to eat and drink diminishes, maintaining adequate nutrition becomes increasingly challenging. Proper nutritional support can decrease some of these difficulties, helping to manage symptoms and improve overall well-being. For many patients, food and drink carry deep emotional and social significance. The act of sharing a meal, even in a limited or symbolic way, can provide comfort and a sense of normalcy. Nutritional support, therefore, plays a critical role in maintaining not just physical health, but also psychological and social-emotional well-being.

Nutritional support in palliative care is not merely about prolonging life, but also about enhancing comfort and quality of life. For many patients, maintaining adequate nutrition can lead

to improvements in energy levels, physical functioning, and overall well-being. These benefits are especially pronounced when nutritional support is part of a comprehensive early palliative care approach that addresses the patient's physical, emotional, and psychological needs. Effective nutritional support can help manage symptoms such as fatigue, muscle wasting, and loss of strength, which are common in advanced illness. By improving nutritional status, patients may experience better symptom control, reduced hospitalizations, and a more stable course of illness. [7] Furthermore, nutritional interventions can play a vital role in preserving a sense of normalcy and dignity for patients, allowing them to maintain their usual eating patterns and enjoy meals with loved ones, even as they approach the end of life.

The decision to provide artificial nutrition and hydration is particularly complex. While artificial feeding can help maintain nutrition when oral intake is no longer sufficient, it also carries risks and can be burdensome for patients. The decision-making process should involve a multidisciplinary team and be based on the patient's individual needs, preferences, and goals of care. In end-of-life care, the goal of nutritional support shifts from preventing malnutrition to enhancing comfort and quality of life. Aggressive interventions may not be appropriate if they cause discomfort or distress. Instead, the focus should be on providing nutrition that aligns with the patient's symptomatic needs and overall care goals. The impact of nutritional support on quality of life is profound, as it can influence not only physical health, but also emotional and social well-being. By addressing the nutritional needs of palliative care patients in a holistic and individualized manner, healthcare providers can help improve the quality of life for patients during their final stages of life.

4.2 Social Science

- Role of social workers
- Addressing financial and logistical challenges

Social workers play a major role in palliative care by engaging in a broad range of activities that extend beyond traditional psychosocial support to include advocacy, care coordination, and addressing the complex needs of patients and families facing serious illness, death, and processing grief.

On the individual level, social workers must possess competencies like effective communication, cultural competence, and the ability to integrate holistic and transformational social work practices into palliative care settings. These capacities allow social workers to collaborate effectively with medical and paramedical professionals, evaluate and document their own practices, and provide tailored interventions that address the unique needs of each patient and family. For example, social workers' confidence and competence in palliative care practice are closely tied to the palliative care content within their educational courses, highlighting the need for comprehensive training programs that equip social workers with the necessary skills and knowledge. Contextual factors, including the composition of social work educational courses, the

care models in which social workers operate, and the characteristics of their employing organizations, also significantly influence social work involvement in palliative care.

Social workers are uniquely positioned to address the financial and logistical challenges that often accompany palliative care, particularly for underserved populations. Financial difficulties, including the high costs of care, loss of income due to illness, and inadequate insurance coverage can increase the stress and burden on patients and their families. Social workers help navigate these challenges by connecting families to available financial resources, advocating for patient access to necessary services, and assisting with applications for financial aid, insurance benefits, and other forms of support. By decreasing financial stressors, social workers enable patients and families to focus more fully on quality of life and symptom management. Logistical challenges, such as transportation to medical appointments, obtaining medical supplies, and managing complex care schedules are also critical areas where social workers provide support. [8] They collaborate with community resources to secure transportation, home care services, and other practical needs that can significantly impact a patient's ability to receive palliative care. Social workers also coordinate care across various service providers, ensuring that all aspects of a patient's care plan are communicated and implemented effectively.

The involvement of social workers in palliative care goes beyond individual patient interactions; it also includes broader community work, advocacy, and the pursuit of social justice. As advocates, social workers strive to address systemic barriers that hinder access to palliative care, particularly in marginalized communities. Their role in community education and outreach is vital in raising awareness about palliative care options and promoting equitable access to services. Future research should explore the contextual factors that influence the involvement of social workers in palliative care, particularly how different care models, organizational structures, and policy environments affect their practice. There is also a need to expand research beyond English-speaking countries to understand the diverse roles of social workers in palliative care globally. Social workers must continue to elevate their influence in the broader practice field of death, dying, and bereavement, engaging in community work, advocacy, and the pursuit of social justice to maximize their contributions to palliative care.

5. Ethical Considerations

5.1 Ethical /Legal Decision-making

- Common ethical dilemmas in palliative care
- Guidelines and frameworks for ethical decisions

End-of-life care often presents challenging ethical choices for healthcare professionals, patients, and families to make, particularly when deciding whether to pursue life-prolonging interventions or prioritize comfort and quality of life of the patient. Some key decisions involve

whether to administer cardiopulmonary resuscitation (CPR), use mechanical ventilation (MV), provide artificial nutrition and hydration (ANH), and consider terminal sedation. [9] Each intervention comes with its own set of ethical complexities, often requiring careful consideration of the patient's autonomy, quality of life, and the potential benefits compared to the harms of the treatment.

CPR, while lifesaving in certain situations, may be inappropriate or even harmful for terminally ill patients, often leading to complications that can worsen the patient's quality of life. The decision to perform CPR should consider the patient's preferences, the likelihood of success, and potential risks. In many cases, patients may opt for a Do Not Resuscitate (DNR) order, but conflicts can arise if family members push for resuscitation despite the patient's wants. Ethical principles emphasize that the patient's autonomy must be respected if they are competent to make decisions, revealing the importance of clear communication between healthcare providers, patients, and families.

Similarly, the use of mechanical ventilation and other supportive therapies like extracorporeal membrane oxygenation (ECMO) or mechanical circulatory support (MCS) in end-of-life care is often aimed at decreasing symptoms rather than prolonging life. Decisions to withdraw or withhold these interventions should be guided by the patient's goals, quality of life, and the perceived benefits of continued support. Discussions about the end of these life-sustaining treatments are sensitive and require healthcare providers to balance ethical principles, including respect for the patient's autonomy and the ethical duty to avoid unnecessary suffering.

Artificial nutrition and hydration (ANH) is another critical consideration in end-of-life care, as it fulfills basic human needs but may also pose significant risks, such as aspiration pneumonia and gastrointestinal discomfort. While ANH can improve the quality of life for certain patients, its benefits are often limited and controversial in cases of advanced illness. Decisions regarding ANH must involve a thorough discussion of potential benefits and harms with the patient or their proxy decision-maker. This process emphasizes the ethical importance of informed consent and shared decision-making in end-of-life care.

Overall, navigating these complex decisions requires healthcare professionals to employ a patient-centered approach, prioritizing clear communication, empathy, and respect for the patient's values and preferences. Effective ethical decision-making in end-of-life care aims to balance medical possibilities with compassionate care, focusing on enhancing the quality of life for patients in their final stages.

5.2 End-of-Life Care Preferences

- Importance of advance directives
- Family and caregiver roles in decision-making

Advance directives play a critical role in ensuring that individuals maintain control over their medical care, even in situations where they are unable to communicate their preferences

due to illness or incapacity. These legal documents provide a written outline of a patient's wishes regarding life-sustaining treatments, medical interventions, and end-of-life care. An advance directive often includes a living will and a durable power of attorney for healthcare (DPOAHC), [10] allowing patients to designate a healthcare proxy who will make decisions on their behalf. By establishing clear preferences, patients reduce the emotional and ethical burden on their families, helping to prevent disagreements and confusion during an already stressful time.

One of the main advantages of advance directives is that they ensure that medical interventions align with the patient's values and desires. In end-of-life situations, treatments such as resuscitation, mechanical ventilation, and artificial nutrition can prolong life, but may also cause discomfort or limit quality of life. When patients have clear directives, they can opt to ignore aggressive treatments that do not align with their values. This planning empowers individuals to make informed choices about their care, ultimately preserving their autonomy. Families and caregivers often play an essential role in implementing end-of-life care preferences, especially when an individual is unable to advocate for themselves. In cases where advance directives are not available or specific instructions are lacking, family members may be asked to make decisions regarding medical treatments. This can be a challenging and emotional process, as they must balance their own beliefs with what they believe their loved one would want.

When advance directives are in place, they serve as a critical guide for families and caregivers. These documents provide clarity and guidance, reducing the likelihood of conflict among family members about what type of care should be provided. The healthcare proxy, appointed through the DPOAHC, is tasked with advocating for the patient's wishes, ensuring that treatments align with the patient's documented preferences. This role requires careful communication with healthcare teams and, at times, making difficult choices in the face of complex medical situations. Families and caregivers must also navigate the emotional and ethical challenges of end-of-life decision-making. Open discussions with healthcare providers about the patient's condition, prognosis, and treatment options are essential to making informed decisions. [11] Collaborative care teams, including nurses, social workers, and palliative care specialists, can provide valuable support by facilitating these conversations and offering guidance on how to respect the patient's wishes while providing compassionate care. Ultimately, the involvement of families and caregivers in the decision-making process is crucial to ensuring that end-of-life care respects the dignity and values of the patient, even when they are no longer able to express those wishes themselves.

6. Technology and Informatics

6.1 Technological Advancements

- Use of telemedicine and electronic health records
- Benefits and challenges of integrating technology

The integration of telemedicine and electronic health records (EHRs) in palliative care has transformed how care is delivered to patients with serious, life-limiting illnesses. As telemedicine continues to grow in popularity, it is becoming a vital tool for improving access to care, particularly for patients living in rural or underserved areas. The implementation of EHRs, alongside telemedicine, facilitates communication between multidisciplinary teams, and streamlines patient management, enhancing the quality and coordination of care.

Telemedicine offers several benefits for palliative care, particularly for patients who face mobility challenges or live far from healthcare facilities. Many patients, especially those in advanced stages of illness, experience significant symptom burdens, including pain, shortness of breath, and fatigue, which can make traveling to appointments difficult. Telemedicine allows these patients to receive specialized care from the comfort of their homes, avoiding unnecessary hospital visits and reduces the risk of exposure to other illnesses. For example, virtual consultations enable real-time assessment and adjustment of pain management regimens, as well as timely interventions for symptoms like nausea and anxiety. This not only enhances the patient's comfort, but also supports continuity of care without the physical limitations associated with in-person visits. Studies have demonstrated that telemedicine can improve patient satisfaction, symptom control, and overall quality of life, because patients feel more empowered to manage their conditions in familiar surroundings. Furthermore, caregivers can participate in virtual consultations, allowing them to remain actively engaged in care planning and decision-making processes.

Beyond patient care, telemedicine reduces logistical challenges for healthcare providers, including travel time and costs, while optimizing clinical workflows. Interdisciplinary team meetings and consultations with specialists can be conducted remotely, ensuring timely access to expertise across long distances. This is particularly beneficial in rural regions where access to specialized palliative care services is often limited. The use of EHRs in palliative care plays a critical role in ensuring comprehensive, coordinated care for patients with complex medical needs. EHRs allow for real-time documentation and sharing of patient information among members of the multidisciplinary team, including physicians, nurses, social workers, and pharmacists. This enhances communication and reduces the risk of errors in medication management, treatment plans, and symptom monitoring.

In palliative care, where timely adjustments in symptom management are crucial, EHRs provide immediate access to patient histories, medication lists, and care goals, allowing healthcare providers to make informed decisions. For example, if a patient experiences worsening pain or side effects from medication, providers can quickly review the patient's records, adjust treatment, and communicate changes to the entire care team. EHRs also facilitate the integration of advance care planning documents, such as advance directives and do-not-resuscitate (DNR) orders, ensuring that patient preferences are respected. Moreover, EHRs enable efficient monitoring of clinical outcomes and quality indicators in palliative care programs. The ability to track metrics such as symptom relief, hospital readmissions, and patient

satisfaction helps healthcare systems evaluate the effectiveness of care and make data-driven improvements.

Despite the numerous benefits, integrating telemedicine and EHRs into palliative care also presents several challenges. One of the primary barriers is the limited ability to conduct a comprehensive physical examination via telemedicine. While video consultations allow providers to observe patients and guide them through self-assessments, certain aspects of care, such as palpation or auscultation, are challenging to perform remotely. This limitation can make it difficult to diagnose or assess subtle changes in a patient's condition, especially when relying solely on patient-reported symptoms. Another challenge is ensuring equitable access to telemedicine services, particularly in rural or low-income areas where patients may have limited access to reliable internet connections or the necessary technological devices. The digital divide can increase existing healthcare disparities, particularly for older patients who may be unfamiliar with or uncomfortable using technology. Additionally, privacy and data security concerns are heightened when sensitive patient information is transmitted over telehealth platforms. Patients using unsecured networks or shared devices may expose their medical information, leading to potential breaches of confidentiality. Finally, the use of EHRs can pose challenges related to interoperability and provider burden. [12] In many healthcare systems, different providers may use incompatible EHR systems, making it difficult to share patient information seamlessly across care settings. Additionally, the extensive documentation required by EHRs can contribute to provider burnout, with clinicians spending more time on electronic paperwork than on direct patient care.

The integration of telemedicine and EHRs into palliative care has the potential to greatly improve access, coordination, and quality of care for patients with serious illnesses. While telemedicine offers clear benefits, such as reducing travel burdens and enhancing real-time symptom management, the limitations of remote assessments and technology access must be addressed. Likewise, EHRs provide critical tools for information-sharing and decision-making, but require careful management to ensure interoperability and minimize provider burden. As healthcare systems continue to expand the use of technology in palliative care, ongoing research and policy development will be essential in optimizing its implementation and addressing these challenges.

6.2 Patient Monitoring and Data Management

- Tools for symptom tracking and communication
- Case studies on technology improving patient care

Technological advances in palliative care, particularly for older adults receiving home-based services, have transformed patient monitoring and communication. Telepalliative care tools, including videoconferencing, remote symptom monitoring, and telephonic communication, facilitate real-time interactions between clinicians and patients, ensuring that

care is accessible even in the most remote areas. Symptom tracking technologies provide a systematic way for healthcare providers to collect data, identify changes in patients' health status, and intervene promptly when needed.

Remote symptom monitoring systems offer patients the ability to report symptoms such as pain, fatigue, and nausea in real-time, allowing for more personalized care. These tools enhance patient autonomy, as they allow patients to actively participate in their care by documenting their symptoms and triggering immediate clinical responses when necessary. For example, symptom monitoring systems may include automated alerts to clinicians when pain levels exceed a certain threshold, ensuring prompt palliative interventions to prevent unnecessary suffering. Furthermore, these systems can help reduce emergency hospital visits, as symptom exacerbations are managed early through virtual consultations.

Several case studies highlight the benefits of integrating technology into palliative care. A scoping review identified various technological tools used by older adults in home palliative care, finding that most were comfortable with technology despite initial concerns about usability. One such study examined the impact of video conferencing tools on the continuity of care for terminally ill patients living at home. The results showed that even though older adults preferred in-person eye contact, they still felt supported and closely monitored through telepalliative care. Patients reported feeling less isolated and more confident knowing that clinicians could respond to their needs quickly.

Another case study explored the use of wireless devices for remote monitoring of patients' vital signs, allowing healthcare teams to detect subtle changes in health before symptoms worsen. This proactive approach improved the quality of care by enabling clinicians to address issues before they escalated. Patients reported a sense of security and empowerment, knowing that their health was being constantly monitored.

The COVID-19 pandemic further highlighted the importance of telehealth in palliative care. One study found that patients who had access to telehealth tools experienced fewer disruptions in their care, as clinicians could assess symptoms, adjust medications, and provide counseling virtually. This allowed patients to remain at home, avoiding the risks associated with in-person visits during the pandemic. While these technologies improve access to care, certain barriers remain, such as the usability of devices for older adults, internet connectivity issues, and a lack of familiarity with technology. Addressing these challenges through better training, simplified interfaces, and tailored designs will be key to improving patient outcomes.

Technological tools for symptom tracking and communication have significantly improved the quality of palliative care, particularly for older adults in home settings. These systems not only enhance patient autonomy but also improve the efficiency of care delivery. Future research should focus on optimizing these tools for the unique needs of palliative care patients, ensuring that technology serves as a bridge between clinicians and patients in their final days, rather than as a barrier.

7. Physical and Occupational Therapy

7.1 Role of Physical Therapy

- Techniques to improve mobility and reduce pain
- Tailored physical therapy plans based on patient needs

Physical therapy (PT) plays an essential role in palliative care, addressing both the physical and functional dimensions of patients' well-being. As part of a multidisciplinary care team, physical therapists work to improve mobility, reduce pain, and enhance the overall quality of life for individuals with life-limiting conditions such as cancer, neurodegenerative diseases, and chronic respiratory disorders. By focusing on symptom control and rehabilitation, PT interventions are critical for maintaining independence and preserving dignity in patients during the advanced stages of their illnesses. The flexibility in delivering physical therapy, whether in inpatient facilities, hospices, or home care settings, further makes sure that patients receive personalized care suited to their needs.

In palliative care, the primary goal of physical therapy is to manage the physical discomforts associated with illness progression while also addressing functional limitations that arise as a result of declining health. The physical aspects include the management of pain, muscle weakness, respiratory difficulties, and general fatigue, which are common challenges faced by palliative patients. Physical therapists use a variety of methods, including therapeutic exercises, electrical modalities, thermal treatments, and manual therapy, to decrease these symptoms and improve overall physical function. Importantly, these interventions are tailored to each patient's condition, taking into account the stage of the disease and the individual's specific goals, such as pain relief, improved mobility, or enhanced comfort.

One of the primary methods used by physical therapists in palliative care is therapeutic exercise, which includes passive, active, and assisted movements designed to improve muscle strength, flexibility, balance, and endurance. These exercises may be performed on land or in water (hydrotherapy), depending on the patient's condition and preference. In addition to physical benefits, therapeutic exercises can also promote relaxation and decrease anxiety, contributing to the patient's emotional well-being. By improving mobility, therapeutic exercises help patients perform activities of daily living (ADLs), such as feeding, dressing, and bathing easier, significantly enhancing their quality of life.

Another important technique in palliative physical therapy is the use of electrical modalities such as neuromuscular electrical stimulation (NMES) and transcutaneous electrical nerve stimulation (TENS). These techniques are particularly beneficial for patients suffering from chronic pain or muscle weakness, conditions commonly associated with advanced cancer or neurodegenerative diseases. Electrical modalities provide pain relief by stimulating the nervous system and promoting muscle re-education, helping patients maintain a level of physical function that may otherwise decline as their disease progresses.

In addition to electrical modalities, thermal treatments—the application of heat or cold—are widely used in palliative physical therapy to manage pain, reduce inflammation, and alleviate muscle spasms. [13] For instance, heat therapies like hot packs or ultrasound treatments can soothe aching muscles, while cryotherapy or cold compresses can reduce swelling and numb pain. These modalities are often combined with therapeutic exercises to optimize patient comfort and promote physical recovery.

Manual physical therapy is another vital component of palliative care. Hands-on techniques such as myofascial release, trigger point therapy, and joint mobilization are used to relieve musculoskeletal pain and improve mobility. These techniques, when skillfully applied by physical therapists, can help restore functional movement and reduce physical discomfort. For palliative care patients, manual therapy can provide immediate relief from pain and stiffness, allowing them to maintain a level of physical independence and engage in daily activities with less discomfort. For many patients in palliative care, the use of assistive devices is essential in maintaining independence and mobility. Devices such as orthoses, prostheses, walkers, and wheelchairs allow patients with limited mobility to perform daily tasks with greater ease and safety. Physical therapists not only provide these devices but also train patients and caregivers on their proper use, ensuring that both the patient and their family can manage the challenges of declining physical function.

A key aspect of physical therapy in palliative care is the development of tailored treatment plans that address the unique needs of each patient. These individualized plans take into consideration the patient's movement potential, functional limitations, and personal goals. For instance, a patient with advanced cancer might prioritize pain management and maintaining mobility for self-care activities, while someone with a chronic respiratory disorder might benefit from exercises that improve lung function and promote energy conservation. By creating customized plans, physical therapists are able to address both the physical and emotional aspects of patient care, enhancing their quality of life.

Physical therapy in palliative care is not limited to symptom management, but extends to improving patients' overall sense of well-being by promoting independence, reducing caregiver burden, and fostering dignity. The focus on both physical and functional dimensions of care ensures that patients can experience relief from discomfort while remaining as active and engaged as possible in their daily lives. By incorporating these tailored physical therapy interventions, palliative care teams can support patients in achieving their goals, whether they seek comfort, mobility, or simply the ability to spend more time with loved ones. In conclusion, physical therapy plays a critical role in the comprehensive care of palliative patients. Through therapeutic exercises, electrical and thermal modalities, manual therapy, and assistive devices, physical therapists help alleviate pain, improve mobility, and enhance quality of life. By developing tailored treatment plans that address individual needs, physical therapy contributes significantly to the physical and emotional well-being of patients, allowing them to maintain dignity and independence throughout their palliative care journey.

7.2 Occupational Therapy

- Maintaining independence and quality of life
- Adaptive tools and methods for daily living activities

Occupational therapy plays an important role in palliative care, focusing on enhancing the quality of life for individuals facing life-limiting illnesses. Its main objective is to support clients in maintaining as much independence as possible, despite their declining physical or cognitive abilities. By utilizing adaptive tools, personalized care plans, and rehabilitative techniques, occupational therapy aims to help individuals keep their dignity, perform daily activities more effectively, and ultimately improve their overall well-being. As such, it is a fundamental aspect of palliative care that addresses not only physical needs but also emotional and psychological well-being.

A core principle of occupational therapy in palliative care is to enable clients to participate in activities of daily living (ADLs), which include basic tasks such as dressing, eating, personal hygiene, and using the bathroom. The ability to perform these tasks significantly affects an individual's functional status and quality of life. For clients facing life-limiting illnesses, the decline in physical strength, coordination, and cognitive function can make even basic ADLs challenging. This loss of independence can lead to feelings of helplessness, diminished self-worth, and reduced motivation. Therefore, maintaining as much autonomy as possible through adaptive strategies and tools is not merely about preserving physical abilities; it is about upholding the client's sense of identity and dignity during a difficult phase of life.

To address the challenges faced by palliative care clients, occupational therapists use a range of adaptive tools and techniques designed to compensate for declining physical abilities. These adaptations often focus on simplifying tasks, minimizing physical exertion, and ensuring safety. For instance, clients experiencing reduced dexterity or strength can benefit from using assistive devices such as utensils with ergonomic handles, button hooks for clothing, or electric toothbrushes. These modifications allow individuals to continue performing tasks with less effort and discomfort, helping to preserve their independence. Adapting the environment is another crucial aspect of occupational therapy in palliative care. Environmental modifications aim to reduce barriers that may hinder the client's ability to engage in daily tasks. This can involve installing grab bars in bathrooms, using non-slip mats, or arranging furniture to create open spaces for easier movement. For clients with mobility impairments, equipments like walkers, wheelchairs, or shower chairs can provide necessary support to perform daily activities safely and effectively. The goal is to create a living space that accommodates the client's evolving needs and creates an atmosphere where they can continue to participate in essential tasks.

In addition to adaptive tools and environmental modifications, occupational therapy in palliative care may involve training clients in energy conservation techniques to help them manage fatigue, a common symptom in life-limiting illnesses. This approach includes strategies such as pacing, prioritizing tasks, and taking frequent breaks to conserve energy throughout the

day. Clients are encouraged to break down complex tasks into smaller, manageable steps, allowing them to complete activities without becoming overwhelmed. For example, sitting while dressing or organizing items at waist height can reduce physical strain and improve task efficiency. Another key element of occupational therapy in palliative care is the holistic assessment and ongoing evaluation of the client's abilities. [14] Given that the symptoms and functional status of individuals in palliative care can fluctuate, it is crucial to regularly reassess their needs and adapt care plans accordingly. Occupational therapists work closely with other healthcare professionals, clients, and their families to establish realistic goals, prioritize tasks, and provide the necessary support for adapting to changes in physical or cognitive function. This collaborative approach helps ensure that the care provided remains flexible and responsive to the client's evolving needs.

While the physical benefits of occupational therapy in palliative care are evident, its impact on emotional and psychological well-being is equally significant. The ability to participate in ADLs, even with assistance or adaptive tools, can offer a sense of accomplishment and purpose. For many clients, this participation in daily life tasks can enhance their mood, reduce anxiety, and provide a sense of normalcy. Occupational therapy helps counter feelings of powerlessness by empowering clients to be active participants in their care. Furthermore, involving family members or caregivers in the occupational therapy process can strengthen the support network around the client. Training family members to assist with daily tasks, use adaptive equipment, or modify the environment fosters a collaborative care approach that can alleviate the emotional burden on the client and their loved ones. This shared responsibility and understanding of the client's needs can create a more supportive and compassionate caregiving environment.

Occupational therapy in palliative care goes beyond addressing physical limitations; it is about preserving a client's autonomy, dignity, and quality of life. Through the use of adaptive tools, environmental modifications, and energy conservation techniques, occupational therapists empower individuals facing life-limiting illnesses to continue participating in meaningful daily activities. This holistic approach not only enhances physical function, but also addresses the emotional and psychological needs of the client, fostering a sense of purpose and well-being. As a dynamic and client-centered component of palliative care, occupational therapy helps ensure that the care provided is tailored to the unique experiences and aspirations of each individual.

8. Conclusion

- Recapitulation of the critical roles of technical sciences in palliative care
- Future directions for research and practice + best practices
- The potential impact of emerging technologies and interdisciplinary collaboration

Palliative care plays an essential role in improving the quality of life for patients with life-limiting illnesses and their families, often relying on technical sciences to drive progress in

practice. It is a field deeply influenced by advances in technology, which continue to shape communication, decision-making, and symptom management. However, despite progress, challenges remain, especially in improving the patient experience at the individual level. Communication is fundamental to effective palliative care, influencing treatment choices, patient compliance, and family dynamics. However, many healthcare providers encounter difficulties when discussing terminal diagnoses, the transition to palliative care, or complex family relationships. To address these gaps, further research is needed to understand the nuances of communication in palliative settings.

Key areas of focus should include studying the impact of communication on patient outcomes, examining how practitioners can better navigate challenging conversations, and developing family-centered communication interventions. Additionally, research should explore how healthcare providers can address denial or resistance from patients and families, providing them with tools to foster open, honest dialogue while maintaining empathy. Effective decision-making in palliative care is crucial for ensuring that patients' wishes are respected and that families are supported during emotionally taxing times. However, family dynamics and surrogate decision-making can complicate this process, leading to tension and delays in care. Research should aim to develop models that facilitate shared decision-making while recognizing the diverse roles that family members play in these conversations. Ethical considerations regarding surrogate decision-making also warrant further exploration, especially in situations where a patient's capacity to make decisions may be diminished due to illness.

While pain management is a well-established area in palliative care, less attention has been given to other common symptoms such as fatigue and sleep disturbances. Research should prioritize non-cancer patients and explore non-pharmacological treatments, including complementary therapies and self-management programs. Additionally, the development of tailored interventions for patients with advanced chronic illnesses can improve overall symptom management and enhance quality of life. Moving forward, palliative care research should aim to integrate technological advancements and promote interdisciplinary collaboration. Telehealth and wearable devices, for instance, have the potential to improve symptom monitoring and communication between patients, families, and care teams. Furthermore, "bedside to bench" research approaches can help translate effective clinical practices into evidence-based interventions, ensuring that best practices are continuously refined.

Technical sciences play a critical role in advancing palliative care, especially when it comes to communication, decision-making, and symptom management. Future research should aim to refine individual-level practices by studying how communication can be improved in palliative settings, addressing decision-making complexities, and broadening symptom management beyond pain relief. Emerging technologies and interdisciplinary collaboration offer promising avenues for enhancing care, ultimately ensuring that palliative care remains patient-centered and responsive to evolving needs. The integration of these innovations will drive the future of palliative care, shaping best practices and ensuring more tailored, effective interventions for patients facing life-limiting illnesses.

Works Cited

- [1] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5859921/>
- [2] <https://www.pace.edu/palliative-care/learning-center/pain>
<https://www.ncbi.nlm.nih.gov/books/NBK532303/> [3]
<https://www.webmd.com/palliative-care/end-of-life-coping-with-anxiety-and-depression>
- [4] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6388408/>
- [5] <https://www.nursinginpractice.com/clinical/nutritional-management-in-palliative-care/>
- [6] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7911232/>
- [7] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8637690/>
- [8] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7958189/>
- [9] <https://www.nia.nih.gov/health/advance-care-planning/advance-care-planning-and-health-care-decisions-tips-caregivers-and>
- [10] <https://www.cheshiremed.org/health-wellness/article/importance-advance-care-planning#:~:text=Completing%20an%20advance%20directive%20will,yourself%20while%20you%20are%20able.>
- [11] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11054863/>
- [12] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3012236/>
- [13] <https://www.ausmed.com/learn/articles/activities-of-daily-living>
- [14] <https://pmc.ncbi.nlm.nih.gov/articles/PMC4605199/>

Tariffs and Their Unexpected Consequences: The Impact of US-China Trade Restrictions on the US and Third-Party Countries By Ian Wong

Abstract

The trade disputes between the United States (US) and China during the 2017-2021 Trump presidential administration sparked further debate on the impact of trade wars for the countries directly involved. However, there is a weaker understanding of how these trade disputes affect other countries as trade redirections are employed to evade trade restrictions. Therefore, this paper seeks to explore the following research question: *what have been the effects of Trump's 2018 tariffs on the US and third-party countries impacted by the shift in trade flows?* This paper employs quantitative research methods to investigate the effects of these tariffs by analyzing trade flow data from the US, China, and third-party countries engaged as trade intermediaries for China from 2018 to 2021. This research suggests that China relied significantly on third-party countries for exports to the US, these third-party countries benefited financially from acting as trade intermediaries, and US consumers experienced higher prices due to the tariffs and their side effects. Consequently, trade policymakers should be aware of these direct and indirect impacts when deciding to implement tariffs and determining how to structure trade policies to prevent the usage of third-party countries as intermediaries.

Introduction

The newly elected president, Donald Trump has expressed interest in increased tariffs and protectionist measures against China. He has stated that he will implement tariffs of 60% on Chinese goods if he is re-elected (Politico, 2024). Another possibility is to repeal China's "Permanent Normal Trade Relations" status, causing it to be subject to tariffs of around 40% (Yao, 2000). This would be a continuation of Trump's highly protectionist stance against China during his first presidential term, where he introduced tariffs on thousands of products imported from China valued at approximately \$380 Billion, resulting in a total tax increase of nearly \$80 Billion (York, 2024). To that end, it is important to explore and understand the potential effects and outcomes of these proposed policies on different stakeholders.

According to economic theory, when the US imposes a tariff on China, imports flowing into the US from China which are affected by the tariffs will be substituted by less productive exporters from the rest of the world and in the US causing an increase in prices. However, there may be a possibility that this theoretical shift in trade flow is not fully realized. This could be due to either a misreporting of trade flows or trade redirection (transshipments) where goods are shipped to an intermediate destination before reaching their final destination (Haberkm et al., 2024). This would mean that the significance of the tariff on Chinese goods was less than expected, possibly allowing third-party countries to benefit in the case of transshipments. In all cases, the effect of the tariffs whether according to theory or not is inflationary. Domestic and foreign producers might not have a competitive or absolute advantage over Chinese producers, meaning that these producers incur higher costs of production. These introduce costs into the

market through opportunity costs and an increase in the actual cost of the goods. Trade redirection will also increase prices due to increased shipping and handling costs, albeit not as much as the tariffs. With these increased prices there would be a marginal group of consumers who are excluded from the purchase of the goods (Fajgelbaum & Khandelwal, 2021).

This paper seeks to explore the following research question: *what have been the effects of Trump's 2018 tariffs on China and third-party countries impacted by the shift in trade flows?*

This paper also identifies reasonable patterns in trade shifts that might hold given a further increase in tariffs on China by the US. It also explores the incidence of these changes in the flow of goods and how they might affect stakeholders such as third-party intermediaries or US consumers. This is done broadly by comparing trade reporting between the US and China and the trade flows of tariff and non-tariff goods.

First, this paper will present a timeline of the trade war, highlighting key moments where significant changes in trade flows could be expected and giving context to the different tranches in which the tariffs were introduced. This paper will then discuss the methods by which the data used was procured and processed, followed by an analysis and interpretation of the data, where the potential factors causing certain observed trends will be discussed. Lastly, certain conclusions about the implications of these trends on different stakeholders will be drawn.

To better understand the impacts of the tariffs imposed by the US on China, it is important to review the sequence of events which was the US-China trade war.

Background

The US-China trade war rose in prominence in 2018 due to escalating tensions over China's trade practices, in particular accusations of unfair practices and intellectual property theft (BBC, 2020). Then, President Trump, aiming to reduce the significant trade deficit and protect American jobs, imposed the first wave of tariffs.

The US introduced its tariffs on China in four waves (tranches). Tranches 1 to 3 were announced and introduced in full effect. While tranche 4 was only partially enacted, splitting it into tranche 4A (carried out) and tranche 4B (proposed). Below is a general timeline of the trade dispute between the US and China, with the dates and details of the different tranches bolded, as well as, a general description of the composition of the tariffs.

Timeline of the Tariffs

The timeline below was taken from the South China Morning Post (SCMP, 2021):

Date	Event
06/07/2018 (Tranche 1)	US-China trade war begins as US imposes 25 percent tariffs on US\$34 billion worth of Chinese imports
06/07/2018	China retaliates by imposing 25 percent tariffs on 545 goods originating from the US worth US\$34 billion
23/08/2018 (Tranche 2)	Washington imposes 25 percent tariffs on a further US\$16 billion worth of Chinese goods

23/08/2018	China responds by applying 25 percent tariffs on US\$16 billion worth of US goods
24/09/2018 (Tranche 3)	US places 10 percent tariffs on US\$200 billion worth of Chinese imports
24/09/2018	China responds by placing customs duties on US\$60 billion worth of US goods
01/12/2018	Xi Jinping and US counterpart Donald Trump call a truce in the trade war at the G20 summit in Argentina
10/05/2019	After trade negotiations break down, the US increases tariffs on US\$200 billion worth of Chinese goods, from 10 to 25 percent
15/05/2019	US Department of Commerce announces the addition of Huawei to its “entity list”
31/05/2019	China announces plans to establish its own “unreliable entity list”
01/06/2019	China increases tariffs on US\$60 billion worth of US products
29/06/2019	Xi Jinping and Donald Trump again agreed to a trade war truce, this time at the G20 summit in Japan
05/08/2019	The US designates China as a “currency manipulator”
13/08/2019 (Tranche 4a & 4b)	US announces that various planned levies on US\$455 billion worth of Chinese products have either been delayed or removed
23/08/2019	China announces planned tariffs of 5 and 10 percent on US\$75 billion worth of US goods
01/09/2019 (Tranche 4a)	US tariffs on more than US\$125 billion worth of Chinese imports begin as expected
11/09/2019 (Tranche 4b)	US agrees to briefly delay new tariffs on US\$250 billion worth of Chinese goods
11/10/2019	US announces that it will delay a planned tariff increase of 25 to 30 percent on US\$250 billion worth of Chinese goods
15/01/2020	China and the US sign the phase-one trade deal
14/02/2020	China halves additional tariffs on US\$75 billion worth of American products imposed in 2019
12/05/2020	China announces a second batch of trade-war-tariff exemptions covering 79 American products
14/05/2020	China allows imports of barley and blueberries from the US
01/09/2020	Dozens of US imports from China are granted short extensions to previous tariff exemptions
14/09/2020	US customs agency issues “withhold release orders” banning cotton, apparel, hair products, and computer parts from four Xinjiang companies
15/09/2020	China decides to exempt additional tariffs on a batch of 16 US products for another year
02/12/2020	The US government says it will begin to block the import of all cotton products made by the Xinjiang Production and Construction Corps (XPCC)
02/12/2020	US President-elect Joe Biden tells The New York Times he will not make any “immediate moves” to lift trade war tariffs
18/02/2021	US Treasury Secretary Janet Yellen says that tariffs on China will be “kept in place”

Table 1. Timeline of the US-China trade war (2018 - 2021), SCMP, 2021

Tranche Composition

The general breakdown of which goods made up each tranche is as follows.

Total Imports	Raw materials	Intermediate goods	Consumer goods	Capital Goods
Tranche 1	0.00%	3.53%	9.18%	87.29%
Tranche 2	0.00%	20.39%	5.25%	74.36%
Tranche 3	0.54%	8.91%	42.85%	47.70%
Tranche 4A	0.27%	4.94%	73.65%	20.99%
Tranche 4B	0.00%	2.64%	39.38%	57.98%

Table 2. Tranche composition of good types

Data and Methodology

To investigate the diversion of goods the discrepancies between Chinese exports to the US and US imports from China were investigated. If there was indeed a discrepancy in the reporting of these figures, it would suggest that some external forces are at play in influencing the reporting of these figures. Furthermore, if there are changes in the reporting gap, it could indicate changes to and influences on these external forces which could be attributed back to the US-China tariffs. Data from both sides of the trade flow were used to facilitate this investigation.

a. Data from the US Perspective

The first data source is the United States International Trade Commission (USITC), a US government website. The USITC Data Web attains data from “automated forms and reports filed initially with the U.S. Customs Service or, in some cases, directly with the Census Bureau, for virtually all shipments entering (imports) the United States” (*DataWeb*, 2024). This data is from official US sources and represents how the US views and perceives its trade flows.

Using raw import data from the USITC, the yearly dollar amount of US imports originating from different countries (China, Singapore, etc) for the individual (Harmonized Tariff Schedule - 8) HTS-8 codes, for example (8470.50.00 - Cash Registers) were attained. This categorization facilitates a tagging of the individual goods to their corresponding tranche or lack of inclusion in a tranche.

b. Data from the Chinese perspective directly from Chinese reporters

The second data source is the General Admissions Council of the People’s Republic of China’s customs statistics website (海关统计数据查询平台, 2024). This source provided an aggregate value of China’s monthly exports to trading partners, specifically, China's exports to the US.

This data being directly from an official Chinese source is more representative of how China perceives its trade flows, however, it has its limitations. The data was not categorized into HS codes and thus could not be used if the goods needed to be sorted into their respective tranches. To compensate for this, Chinese data from a third party - UNComtrade was sourced.

c. Data from the Chinese perspective through a UN commission

The third data source is the United Nations Department Of Economic and Social Affairs statistics website (*UN Comtrade*, 2024). UN Comtrade ensures its data is accurate and reliable by relying on official sources. Typically, this involves working with a single agency in each country, such as a customs department, national statistics office, or a relevant ministry. UNComtrade collaborates directly with these official bodies to gather the data. Although regional organizations may sometimes be used as secondary sources, UNComtrade prioritizes data that comes directly from the responsible national authority. This approach guarantees that the data aligns with each country's official reporting and methods (“The (Essential) Role of UN’s Comtrade in Trade Data”).

From UN Comtrade, Chinese export data to different countries by HTS code on an annual basis (similar to the first data source) was obtained. Using this data which includes the HS codes, the same tagging as mentioned previously was able to be carried out, categorizing the goods into their different tranches.

Methodology

This paper utilizes a quantitative research methodology drawing on data obtained in Excel format from all three sources (USITC, the General Admissions Council of the People’s Republic of China’s customs statistics website, and UN Comtrade). Data from USITC and UN Comtrade were organized by HTS-8 and HTS-6 codes respectively. For easy comparison between the two sources, the data from the USITC was aggregated into the HTS-6 level. Using a tagging found on the USITC website that tagged affected HTS-6 codes to their corresponding tranche, all HTS-6 codes in the raw data could be sorted into their respective tranche (or lack thereof). This facilitated the aggregation of goods into two categories: tariffed (tranches 1 to 4a) and non-tariffed (tranche 4b and goods not included in tranches). With this tagging, individual third-party trading partners could be investigated as players in the potential trade redirection, investigating how trade flows changed between these countries and China and the US.

Analysis

In this section, the findings from the processing of the raw data will be discussed, it aims to lay out plainly the findings from the data and provide an interpretation of the trends and patterns observed.

Reporting Gap

The figure below was obtained from data sources 1 and 2, the blue line represents US-reported imports from China, and the orange line represents Chinese-reported exports to the US—both are plotted monthly and are on a 12-month moving average. The gray area represents the gap between these two lines, which can also be called the reporting gap. The figure gives us a general view of the change in US-China trade and how the reporting gap changed over time.

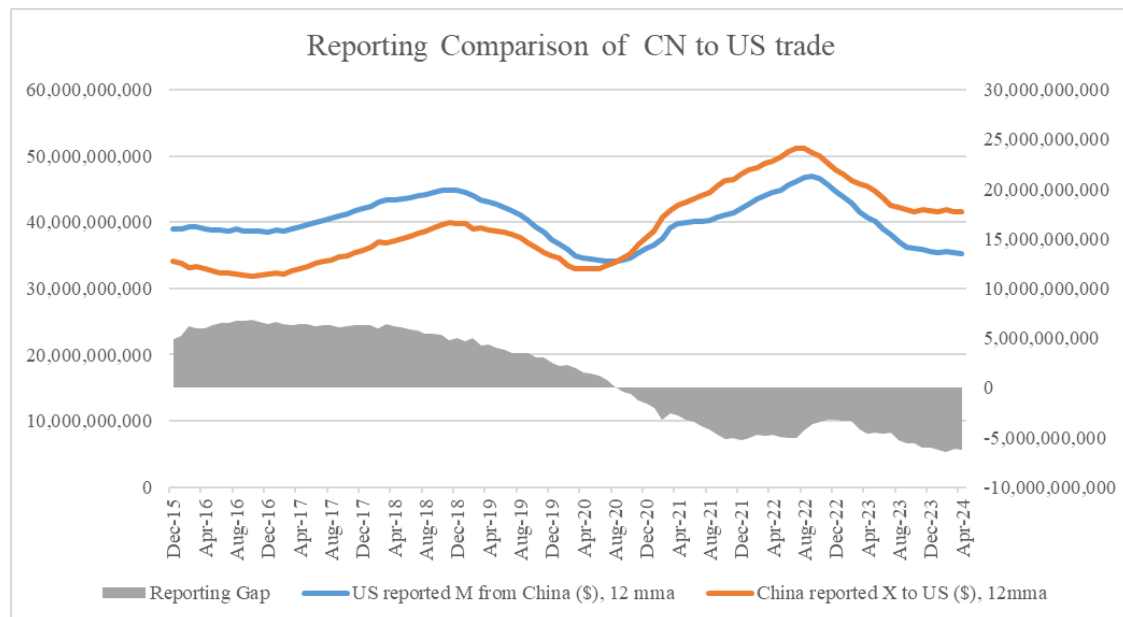


Figure 1. *Change in Reporting Gap between the US and China*

Trade originating from China and entering the US dropped sharply following the introduction of the tariffs around August of 2018, as expected. However, when viewing the actual value of trade between China and the US, there are discrepancies in reporting value depending on which trading partner the data is derived from. There are a few reasons for a gap in the reporting of trade: transshipments/re-exports, trade fraud, and a difference in the method of good valuation.

A reporting gap caused by a difference in valuation methods should result in a relatively constant gap, if all else is constant, assuming that this accounts for the gap in reporting between China and the US before 2018. Unlike a difference in valuation methods, transshipments/re-exports and trade fraud are more reactionary, meaning their incidence can change as a reaction to certain events, like the introduction of Trump’s 2018 tariffs.

These tariffs resulted in an extra cost for US importers. Higher costs raise the incentive for US importers to under-report the value of their imports, this is compounded with possible trade diversion which negatively misrepresents the value of Chinese imports from the US perspective. The incidence of such an occurrence is exemplified not by a reduction of the reporting gap but a total reversal, where the value of trade fraud and transshipments has far overcome the initial difference due to the disparity in valuation methods.

To attempt to gain a better insight into the relationship between the tariffs and the change in the reporting gap, a more in-depth study comparing goods that were tariffed and goods that were not was conducted. The figures below plot the change in reporting of trade similar to Figure 1 but separated into goods that were tariffed and goods that were not.

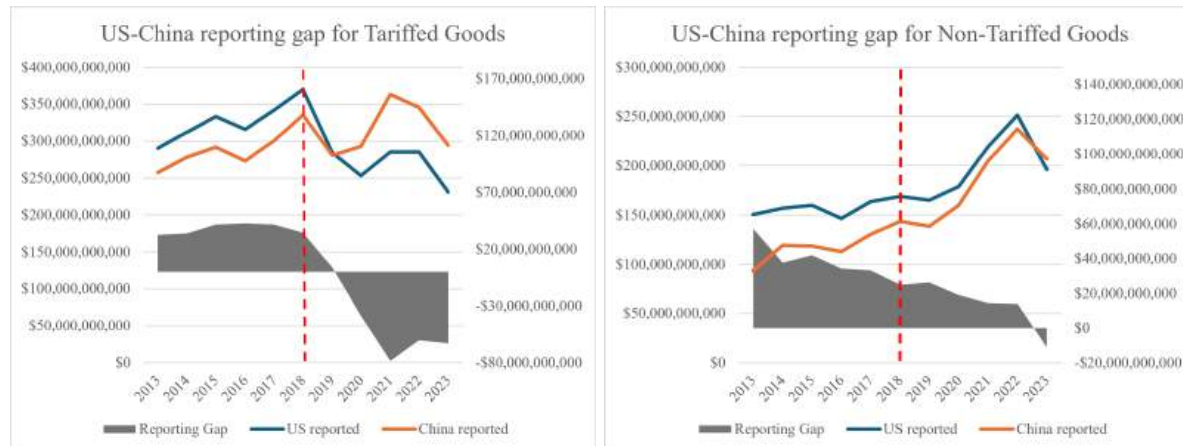


Figure 2. *US-China Trade Reporting Gap for Tariffed and Non-Tariffed Goods*

From the figure, a large shift in the reporting gap for tariffed goods can be observed. The reporting gap shifted from around \$190 billion in 2017 to around - \$50 billion in 2020, a change of - \$240 billion over the years the tariffs were imposed. Conversely, for the non-tariffed goods, only a slight shift of about -\$10 billion from 2017 to 2020 can be seen. This is relatively insignificant when compared to the shift in reporting gap for tariffed goods which is greater by a factor of 24. This serves to emphasize the direct relationship between the introduction of the tariffs and a change in the reporting gap. Validating our claim that trade redirection and/or trade fraud are occurring at a larger scale as a result of the US-China tariffs.

Trade Redirection

From the evidence discussed above, there are two main explanations, trade redirection where trade is being diverted through other countries, and trade fraud where imports are being intentionally underreported. For the purposes of this investigation, the main focus will be on trade redirection. This is not to say that trade fraud does not occur, only that the data readily available to me better suits the investigation of trade redirection.

Certain countries in Asia were chosen as the main stakeholders in facilitating the redirection of goods from China to the US. The countries were chosen for: 1) their geographical proximity to China, and 2) their openness to Chinese influence. The countries include Cambodia, India, Indonesia, Japan, Korea, Malaysia, Thailand, and Vietnam. The figure below plots the flow of tariffed and non-tariffed goods from China to Asia and Asia to the US.

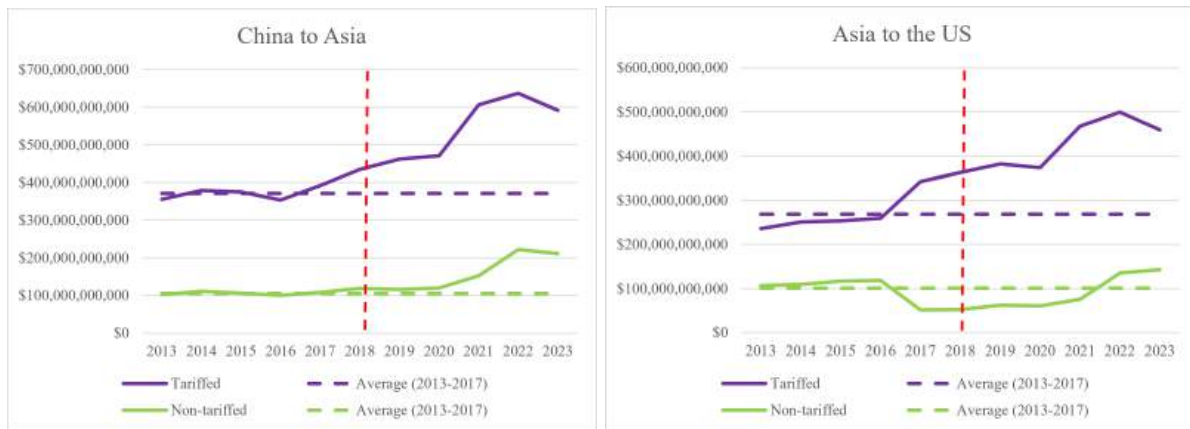


Figure 4. *Change in Trade amounts between China, Asia, and the US for Tariffed and Non-Tariffed Goods*

When observing the change in trade flows between China, Asia, and the US, a significant discrepancy between the increase in flows of tariffed goods as compared to non-tariffed goods can be observed. For trade flows from China to Asia, there is a noticeable increase in the flow of tariffed goods post-2017 while the flow of goods for non-tariffed goods remains relatively the same. This is similar to the trade flow of Asia to the US, where there is a significant increase in the flow of tariffed goods while the flow of non-tariffed goods remains relatively unchanged. This discrepancy can be explained in one of two ways. Firstly, by following normal economic theory: For China to Asia, Chinese exporters will divert a proportion of exports away from the US to other countries where they do not incur a tariff. For Asia to the US, Asian exporters will fill the gap in imports left by the diversion of trade away from the US by China. Secondly, and more interestingly, trade diversion: From China to Asia, Chinese exporters of tariffed goods that are meant for the US will export goods to other countries, likely countries that have a convenient geographical location (countries close to China, i.e. Asia). From Asia to the US, the goods exported to Asia are then re-exported to the US. From the US perspective, these goods are coming from Asia and therefore are not subject to tariffs. In this way, Chinese exporters are still able to “export” their goods to the US, albeit through another country. This would explain the rise in the trade flow of tariffed goods from China to Asia and subsequently Asia to the US.

These two ways of explaining the discrepancy are not mutually exclusive and likely are occurring in tandem. The patterns observed, however interpreted, highlight a fundamental shift in global trade, carrying significant implications for various stakeholders, as discussed in the next section.

Impacts and Consequences

In this section, the general effects of the tariffs based on the evidence and conclusions drawn from previous sections will be discussed.

Third-Party Countries

The imposition of tariffs by the US on Chinese goods has led to significant changes in trade flows and third party countries have likely benefited from this. Countries in Asia, such as Vietnam, Malaysia, and Cambodia, have experienced a sharp rise in trade activities after the introduction of tariffs on China. They have become key intermediaries in transshipment operations. Goods that were once directly exported from China to the US are now being diverted through these third-party countries. This transshipment allows Chinese goods to bypass direct tariffs while still entering the US market.

The middleman economy in these third-party countries gains from their handling fees, logistics, and certain value-added processes. Manufacturing industries, especially those involved in final assembly, could see significant growth. For example, electronics or apparel, where components or unfinished goods may be shipped to neighboring countries for final touches and packaging. The shipping sector in these countries would also have seen an uptick in demand due to increased transshipments, facilitating the movement of goods between China, intermediate countries, and the US.

Although this paper focused largely on countries in Asia, this phenomenon could very well occur in other neighboring countries or countries located along shipping routes. One region of interest would be Central and South America given their proximity to the US.

US Consumers and Producers

US consumers may initially benefit from these transshipments as it enables lower prices on goods that would otherwise face steep tariffs if directly imported. However, the presence of tariffs still introduces inefficiencies into the market, which manifests as a deadweight loss.

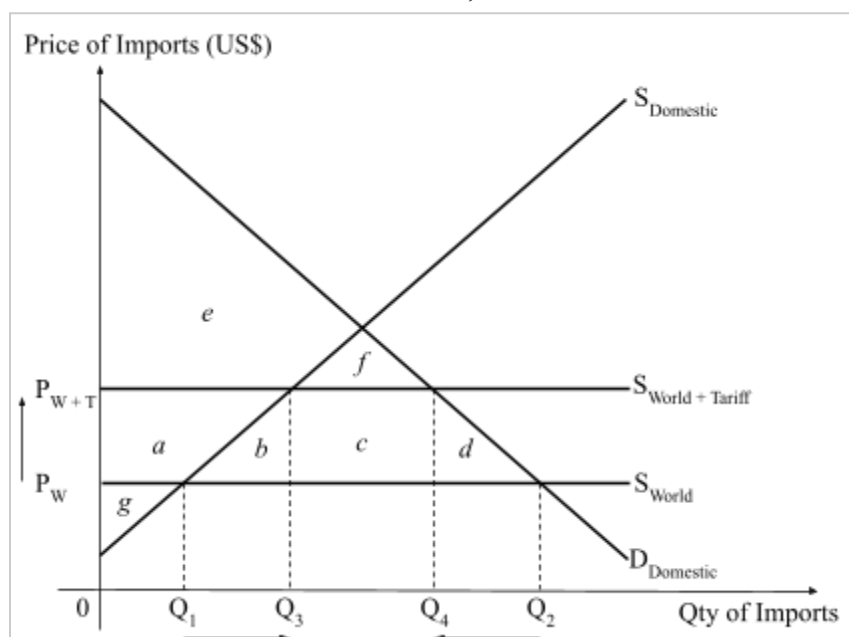


Figure 5. Tariff diagram for the US import of Chinese goods

This model assumes that the US only produces a fraction of the global steel supply, accepts world price (P_w) under free trade (before the US-China tariffs) and China has an absolute advantage in the production of steel. Before the tariffs, under free trade, the price of steel was at P_w , the quantity supplied by domestic producers was at Q_1 and the quantity demanded by domestic consumers was at Q_2 . The US imports the amount from Q_1 to Q_2 from foreign producers to fill the excess demand ($Q_1 > Q_2$). With the imposition of the tariffs, the domestic price of goods rises from P_w to P_{w+T} (world price + cost of the tariff), the quantity supplied by domestic producers shifts from Q_1 to Q_3 and the quantity demanded by domestic consumers shifts from Q_2 to Q_4 . The US now imports the amount from Q_3 to Q_4 from foreign producers.

US producers gained as their producer surplus rose from area g to area $a + g$. On the other hand, US consumers lost as consumer surplus fell from area $a + b + c + d + e + f$ to area $e + f$. Another thing to note, consumers also include producers who acquire materials from these imported goods to be used to create secondary products. So in some cases, local producers also lost out. On the same note, since this diagram represents an aggregate for all the goods affected by the tariffs, it cannot be said that all the goods experience this exact shift and outcome for stakeholders. For example, a real impact of the tariffs on US producers was the loss of jobs in many industries due to the rising costs of certain imported goods and materials.

Furthermore, there are deadweight loss areas b and d , area b represents the loss in consumer welfare as a proportion of US consumers are priced out of buying the goods due to higher costs, and area d indicates the inefficiency when switching from the lower-cost Chinese producer to a higher-cost alternative producer.

The figure below shows the change in the exports of tariffed and non-tariffed goods from the US to the rest of the world. It can give us a good indication of the effects of the tariffs on the US domestic consumer market.

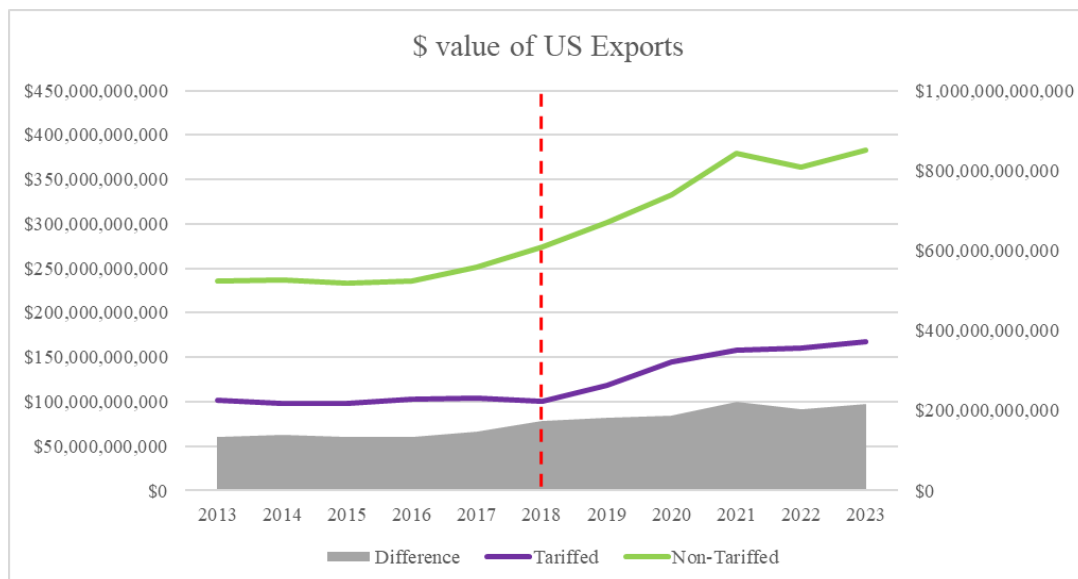


Figure 6. Change in US exports of tariffed and non-tariffed goods

Following the imposition of the tariffs in 2018, US exports generally are observed to rise. Non-tariffed goods, however, experience a greater increase. This can be visualized as the difference between the value of exported non-tariffed goods and the value of tariffed goods increased by almost \$200 billion from 2018 to 2021. This fall in the proportion of tariffed goods in US exports indicates that more tariffed goods are being consumed domestically. This confirms our theoretical understanding that US consumers are consuming more domestic goods and that the prices of said goods are rising since US producers choose to sell a higher proportion of their goods domestically where they can earn a larger profit than in foreign markets.

Future Impacts

Taking our findings of the effects of the tariffs into consideration, it is reasonable to predict that the manufacturing and shipping sectors of third-party countries will experience continued growth, given the future US administration follows through with their promises of increased tariffs.

US consumers are likely to experience persistently high prices on Chinese imports, their supplements, and related secondary products. The demand for supplemental goods (non-Chinese goods) will rise, causing prices to experience upward pressure. The cost to produce Secondary products will rise and to retain profit margins, the cost will be transmitted through producers to consumers in the form of increased prices.

Conclusion

In conclusion, the analysis of the US-China trade war and the subsequent changes in the flows of trade between these two countries reveals an interesting dynamic in the global trade environment. By looking at the impact of the 2018 tariffs, this research highlights how third-party countries have become involved as intermediaries. These countries have benefitted by finding new roles and financial benefits from handling redirected trade. However, this trade redirection, while allowing US consumers to experience short-term price relief, costs US consumers in the long run as added handling and shipping costs, market inefficiencies, and the potential of job losses as costs for certain industries rise. This study illustrates the importance of a careful and measured approach to tariffs, one that considers the unintended consequences and side effects that may tamper with the effectiveness of the tariff.

Evaluation

Research Limitations

This investigation had many strengths, such as the reliance on official data from sources like USITC and UN Comtrade, which provided a strong foundation for my analysis of trade flows. By comparing data from both the US and Chinese perspectives, this paper was able to assess discrepancies in trade flow with a degree of accuracy. The standardization of HTS codes

was also helpful for the categorization of goods into tariffed and non-tariffed, facilitating a nuanced analysis of the impacts on both a macro and micro lens.

However, it also did have its fair share of weaknesses and assumptions. Several assumptions were made to properly analyze the data, particularly concerning the incidence of trade redirection and fraud. The increase in trade redirection through third-party countries was assumed to be driven primarily by the imposition of the US tariffs on Chinese goods. This assumption simplifies the often complex dynamics of global trade, which may also be influenced by other factors like changes in global supply chains, shifts in demand, or other policy changes in the US or China. The effect of the tariffs on goods was assumed to be uniform across the different types of goods within the tariffs. For example, the tariff's impact on consumer goods was assumed to mirror its impact on capital goods. In reality, the sensitivity of these different types of goods varies based on their elasticity of demand and supply. Lastly, the assumption was made that the initial reporting gap (- 2017) was due largely to differences in valuation methods alone and not affected by any other factor that might have been affected by the tariffs. Furthermore, the data used also had its weaknesses, while the data used from UN Comtrade and USITC provided accurate figures, there were gaps in the data, particularly in terms of tracing the shipping flow of the goods and quantifying or even providing any information about trade fraud.

Opportunities for Future Research

Future research should look at other regions beyond Asia, such as Central and South America, and Eastern Europe, to see if similar patterns of trade diversion are occurring. This will give a clearer and more robust understanding of the flow of goods from China to the US.

Individual countries' data should also be examined in future research. By examining individual countries' data in more detail, more detailed insights into the effects of the tariffs on the different sectors of the third-party countries could be uncovered. These countries could include, but are not limited to, the countries included in my aggregation of countries in Asia, Canada, and Mexico.

An exploration into how different types of goods (consumer, capital, and intermediate goods) are affected by the tariffs, should also be conducted. It could provide a clearer picture of the economic consequences for different sectors within the US and third-party countries. Examining intermediate goods could also lead to clearer insight as to which goods in particular could be shipped as unfinished products to third-party countries where they would be given the finishing touches.

Works Cited

- BBC, N. (2020, January 16). A quick guide to the US-China trade war. Bbc.com; BBC News.
<https://www.bbc.com/news/business-45899310>
- DataWeb. (2024). Usitc.gov; USITC DataWeb. <https://dataweb.usitc.gov/>
- Fajgelbaum, P., & Khandelwal, A. (2021). The Economic Impacts of the US-China Trade War *. National Bureau of Economic Research, 14(29315).
<https://doi.org/10.1146/annurev-economics-051420-110410>
- Haberkorn, F., Hoang, T., Lewis, G., Mix, C., & Moore, D. (2024, April 12). Global trade patterns in the wake of the 2018-2019 U.S.-China tariff hikes. Federalreserve.gov.
<https://www.federalreserve.gov/econres/notes/feds-notes/global-trade-patterns-in-the-wake-of-the-2018-2019-u-s-china-tariff-hikes-20240412.html>
- Politico. (2024, February 6). Trump praises Xi but warns China could face heavy tariffs if elected. YouTube. <https://youtu.be/WAg99owBB6w?si=Yd-dkdbX838FQbRF>
- SCMP. (2021, August 29). South China Morning Post.
<https://www.scmp.com/economy/china-economy/article/3146489/us-china-trade-war-timeline-key-dates-and-events-july-2018>
- Shunli, Yao. (2000). US Permanent Normal Trade Relations with China: What is at Stake? -- A Global CGE Analysis *. Social Science Research Network, doi: 10.2139/SSRN.243550
- “The (Essential) Role of UN’s Comtrade in Trade Data.” World Bank Blogs, 2024,
blogs.worldbank.org/en/opendata/the--essential--role-of-un-s-comtrade-in-trade-data.
- UN Comtrade. (2024). Un.org. <https://comtradeplus.un.org/>
- York, E. (2024, July 16). Tariff Tracker: Tracking the Economic Impact of Tariffs. Tax Foundation.
<https://taxfoundation.org/research/all/federal/trump-tariffs-biden-tariffs/#:~:text=The%20Trump%20administration%20imposed%20several,increase%20of%20nearly%20%2480%20billion>
- 海关统计数据查询平台. (2024). Customs.gov.cn. <http://stats.customs.gov.cn/indexEn>