Psychosocial Student Adaptations to COVID-19 in Higher Education: A Mixed-Methods Approach to COVID-19 through the Theoretical Framework of Anomie

Grace Johnson, Butler University
Mentor: Ashley Hutson

Punishment versus Hope: Perceptions of Social Media Portrayals of Restorative Justice

Emily E. Leonard, Butler University
Mentor: Ashley Hutson

Infiltrating the Boys’ Club: The Portrayal of Women in Popular Crime Dramas

Emma D. Bergquist, Butler University
Mentor: Stephen Barnard

Populism: A Comparative Study of the Rise of the Alt-Right in Germany and the United States

Kate O. Tobin, Butler University
Mentor: Melissa Etzler

Stream Gauging: Investigating the Flow Rate and Residence Time in Laguna Bacalar, Mexico

Elizabeth J. Kowalczyk, Middle Tennessee State University
Mentor: Timothy Grundl, Hector Hernández
Modeling Employee Burnout and Stress Levels Using Generalized Linear Models
Madeline S. Neely, Butler University
Mentor: Rasitha R. Jayasekare

Enhanced Breast Cancer Tumor Classification Using MobileNetV2: A Detailed Exploration on Image Intensity, Error Mitigation, and Streamlit-Driven Real-Time Deployment
Aaditya Surya, University of Southern California
Aditya Keshary Shah, Rochester Institute of Technology
Subash Tarum Sasikumar, State University of New York Plattsburgh
Jarnell Kabore, Northeastern University
Mentor: Swapnil Sneham

Opening an Existential Window into Wuthering Heights and The Metamorphosis
Sarah H. Lognion, McNeese State University
Mentor: Joseph Tuminello

The Influence of Specific Diagnosis and Trait Self-Relevance on Positive and Negative Perceptions of Individuals with Mental Illness
Sabrina L. Stock, Butler University
Aedan S. Levy, Butler University
Mentor: Tara Lineweaver

Comparing the Visuospatial Abilities of Alzheimer’s Patients and Huntington’s Patients
Alysa Aralis, Butler University
Emma Bauman, Butler University
Mentor: Tara Lineweaver
Discrimination by Association: The Role of Race and Exoneration in Hiring Decisions

Emma D. Bergquist, Butler University
Mentor: Fabian Alceste

Work-Life Balance in a Medical Unit of U.S. Army Reservists: An Exploratory Study

Elizabeth A. Kleitz, Truman State University
Madelyn Kuhn, Truman State University
Mentor: Carol Cox

Effects of Dissolved Oxygen on Northern Leopard Frog (Lithobates pipiens) Gastrocnemius Muscle Recruitment and Fatigue

Eva L. Gates, Penn State Altoona
Joseph P. Levri, Penn State Altoona
Mentor: Lara LaDage

The Roycean Communal Ideal in T. S. Eliot’s “The Love Song of J. Alfred Prufrock”

Micah Gill, Harding University
Mentor: Nick Boone

Yayoi Kusama’s Depictions of Mental Health

Anna Isabel Davis, Transylvania University
Mentor: Emily Goodman

Partake or Be Departed: Insularity of Setting within the Dystopian Imagination and The Mark: The Beast Rules the World

Andrew Watkins, Fayetteville State University
Mentor: Dean Swinford
VOLUME 10 REVIEWERS

Julia Angstmann  
Office of Sustainability  
Butler University

Steve Barnard  
Sociology and Criminology  
Butler University

Chad Bauman  
Philosophy and Religion  
Butler University

Kelly van Busum  
Computer Science and Software Engineering,  
Butler University

Vitor Dias  
Sociology and Criminology  
Butler University

Heliene Flanzbaum  
English, Butler University

Ryan Flessner  
College of Education,  
Butler University

Lee Garver  
English, Butler University

Samuel Gurevitz  
Health Sciences, Butler University

Brent Hege  
Philosophy and Religion,  
Butler University

Scott Kascher  
Mathematical Sciences,  
Butler University

Joel Martin  
Psychology, Butler University

Dave Mason  
Political Science, Butler University

Fait Muedini  
International Studies,  
Butler University

Kate Novak  
Sociology and Criminology,  
Butler University

Steve Nyktas  
Art + Design, Butler University

Corey Reed  
Philosophy and Religion  
Butler University

Carol Reeves  
English, Butler University

Alex Roehrkasse  
Sociology and Criminology  
Butler University

Travis Ryan  
Office of the Provost,  
Butler University

Betsy Schmidt  
Physician Assistant,  
Butler University

Alex Roehrkasse  
Sociology and Criminology,  
Butler University
MISSION AND EDITORIAL POLICY

The mission of the Butler Journal of Undergraduate Research is to advance Butler’s commitment to undergraduate research by seeking to publish original, high quality empirical research undertaken by undergraduates from any university. The journal is national in scope, multidisciplinary, and seeks to publish the best empirical and scholarly research available from undergraduate students anywhere. Submission of original, scholarly research articles is open to undergraduates from any accredited college and university.

The publication of the Butler Journal of Undergraduate Research is designed to build upon and strengthen Butler’s commitment to quality undergraduate research by providing an outlet for the publication of outstanding undergraduate scholarship. BJUR recognizes and supports diverse disciplines, perspectives and methods across the social sciences, natural and life sciences, the arts and humanities. The Journal publishes papers that represent original research undertaken by a student or students during their undergraduate career.

BJUR is committed to intellectual integrity, rigorous standards of scholarship, and rational and civil discourse. It seeks to foster open and critical inquiry that privileges no particular disciplinary standpoint while operating within the limits of a standard of scientific, scholarly and empirical discourse committed to the pursuit of knowledge and truth. While every effort is made to publish issues that represent a fair balance of scholarship across the entire spectrum of undergraduate studies, the Journal’s first priority is always to publish the best undergraduate research available at any given time regardless of disciplinary representation.

POWERED BY

© 2024 Butler Journal of Undergraduate Research
digitalcommons.butler.edu/bjur
PSYCHOSOCIAL STUDENT ADAPTATIONS TO COVID-19 IN HIGHER EDUCATION: A MIXED-METHODS APPROACH TO COVID-19 THROUGH THE THEORETICAL FRAMEWORK OF ANOMIE

GRACE JOHNSON, BUTLER UNIVERSITY
MENTOR: ASHLEY HUTSON

Abstract

The COVID-19 pandemic affected individuals and social institutions because of an abrupt and expansive change to norms and values, which are key characteristics of an anomic state. The purpose of this exploratory and descriptive study was to understand how students enrolled at a residential university perceived educational and societal pressures associated with COVID-19. This research applied a mixed-methods approach and relied upon three phases of research: (1) pilot questionnaire \((n = 54)\), (2) qualitative interviews \((n = 14)\), and (3) quantitative questionnaire \((n = 253)\). Findings suggested similarities in how university students responded to COVID-19 and how communities had responded previously to other forms of natural disasters. Students indicated experiencing worsening mental health and loss of social connection to others within the community. This project concludes with a discussion of the implications of natural disasters, such as COVID-19, for individual and group responses to strain through resilience, coping mechanisms, and adaptation.

Keywords: COVID-19, anomie, natural disasters, mental health, education

The imagery associated with natural disasters—torn roofs, collapsed homes, and piles of debris—often acts as a visual representation of the damage sustained by communities. Less obvious, however, are the social and emotional tolls felt by individuals and communities (Shing, Jaywickreme, and Waugh 2016). Following large-scale changes, both individuals and communities may feel a sense of anomie, a loss of predictability within everyday life (Durkheim, Sennett, and Riley [1897] 2006). Increasing group and individual disconnection, uncertainty, and anxiety are felt because of the loss of guiding social behaviors and community connections (Merton 1968; Olsen 1965).

In his study of suicide, Émile Durkheim coined the term *anomie* to describe a time without norms (Durkheim, Sennett, and Riley [1897] 2006). *Anomie* is used to describe a pattern of responses to drastic social change through which individuals lose their connections to established social norms, values, and social supports; therefore at
the social level, the lack of norms makes the community dysfunctional (Olsen 1965). As a concept, anomie is particularly useful for understanding the social and psychological responses to COVID-19 because the rapid social change that accompanied this pandemic disrupted society’s established norms and morals.

During COVID-19, stay-at-home orders, new health and safety regulations, and varying widespread media coverage contributed to dramatic shifts in Americans’ group identities and social norms (De la Sablonnière, Lina, and Cárdenas 2019; Isley et al. 2021; Smith and Gibson 2020; Smith, Livingstone, and Thomas 2019); however, these shifts were not uniformly internalized, and two American groups marked by common responses emerged: those who did not adhere to pandemic safety precautions and those who did. Both groups purported to hold the moral high ground, with one group claiming that pandemic restrictions violated their liberties and the other arguing that these restrictions were necessary to protect society’s most vulnerable populations (Savulescu 2023). After the disruption of norms via the outbreak of COVID-19, rather than seeking cross-group social cohesion, the public appeared to further separate, both physically and socially.

Past research has examined individual and group traumas following natural disasters that resulted in changes to or creation of social norms and that could provide a useful framework to understand the societal response to COVID-19. Given the recency of COVID-19, however, scholars have not fully explored the impact of COVID-19 on social norms. Within higher education, academic and social expectations of students rapidly shifted during the spring semester of 2020 at the outbreak of COVID-19 and have continued to evolve since. This direct upheaval led to isolation from peers, a shift from in-person learning to virtual education, and relative separation from the “normal” daily routines of university students. The term psychosocial adaptation is used in this paper to signal study of the psychological experience and social response to COVID-19 and changing normative behaviors.

The current study examines how university students interpreted changes in their academic and personal lives as a result of COVID-19, including personal appraisals of their psycho-emotional well-being. Using a sequential mixed-methods design, I examine university students’ responses to the pandemic and situate these findings within the broader context of social responses to natural disasters. In the section that follows, I relate the existing findings on natural disasters to the COVID-19 pandemic using sociological and psychological literature. Anomie will be used as a theoretical lens for understanding the COVID-19 disaster from a psychosocial perspective. Findings suggest that the university students encountered worsening mental health as well as loss of social connection to others within the community, similar to other community responses after disasters.
Sociological and Psychological Responses to Disasters

Following traumatic group experiences, such as ecological disasters and pandemics, communities often become more cohesive and resilient through an enhanced sense of group unity, trust, and support among surviving members (Almazan et al. 2018; Maki et al. 2018). Resilience, a physiological stress response, is defined by a rapid return to a baseline level of stress following an individual’s processing of a traumatic experience (Shing, Jaywickreme, and Waugh 2016). Social support, community organizations, personal relationships, and health services can all be utilized after disasters to help individuals and groups find resilience and return to a prior expectation of normalcy (Almazan et al. 2018). Often, the creation of shared social identities and helping behaviors allows the community and individual to rebuild their identities via increased social support (Drury et al. 2015).

A variety of coping strategies is associated with resiliency, including (1) acceptance of positive and negative emotions; (2) positive reappraisal of stressors; (3) feelings of acceptance, adjustment, autonomy, and closure; (4) implementation of hobbies and healthy lifestyle choices; (5) access to basic health services, social support, and religious-based communities; and (6) creation of social identities marked by an increased sense of agency (Almazan et al. 2018; Peter and Jungbauer 2018; Shailaja et al. 2020; Shing, Jaywickreme, and Waugh 2016; Smith, Livingstone, and Thomas 2019). Additionally, those who draw upon established support mechanisms, such as religious communities or prior knowledge and experiences, often have an increased sense of resiliency (Almazan et al. 2018; Fu et al. 2018).

Across multiple studies, researchers have identified growth in the use of mental health services on university campuses before the pandemic occurred and a growth of mental health concerns among students following the onset of COVID-19 (Lipson, Lattie, and Eisenberg 2018; Shailaja et al. 2020). Given these two trends, it is likely that mental health services on university campuses continued to experience a further increased demand for services after March 2020, although it also seems plausible that demand for mental health services could have dissipated when university students physically returned to campus and began socializing more during the Fall 2021 semester (Anderson 2022).

The psychological outcomes of COVID-19 are still uncertain, but researchers have documented an increase in anxiety, mental distress, posttraumatic distress, and anger since the pandemic began (Usher et al. 2020). Furthermore, pandemic-related changes could have heightened an individual’s propensity for stress or exasperated mental health concerns; such changes may necessitate increased reliance on support systems and coping strategies (Esterwood and Saeed 2020). Research published after the onset of COVID-19 has identified a potential relationship between the use of adaptative coping strategies and a reduction in anxiety, depression, and stress, as well as increased access to resources and resilience among students (Becker 2020; Shailaja
et al. 2020). For example, in one of a few recent studies on the subject, Fruehwirth and colleagues (2021) compared student mental health pre- and post-pandemic. Among students, their findings indicated increased anxiety, increased feelings of isolation due to the pandemic, and higher rates of reported depression. Although much remains unknown about university students’ emotional and mental health during the pandemic, it appears likely that many experienced negative changes in their mental health.

Attending to the mental health needs of both individuals and groups is critical following disasters. The present study seeks to understand the patterns of mental health outcomes experienced by university students during the COVID-19 pandemic in relation to anomie caused by COVID-19. After review of relevant literature, I suggest that university students may have suffered mental and physical health changes due to the pandemic. Without exercising coping strategies that may improve resiliency, it is possible that university students may continue to experience negative consequences of pandemic-induced anomie.

Research Design

In the present study, anomie was used as a conceptual framework to understand the adaptative patterns of mental health reactions to COVID-19. This study was designed to identify and describe social patterns and trends related to COVID-19 from the perspective of university students enrolled at Midwestern University.\(^1\) This research described patterns of personal mental health during the pandemic, as well as how students adapted to shifting social norms and values from March 2020 to February 2022. There were two guiding research questions:

R\(_1\): How has COVID-19 impacted university students in relation to their evaluation of personal mental health?

R\(_2\): How have university students adapted to the shift in norms and values in response to COVID-19?

This study occurred in three phases.\(^2\) Given the recency of the COVID-19 pandemic, exploratory research was conducted at the first phase, using a pilot survey. The second phase included both exploratory and descriptive research, using qualitative interviews, and the third phase was solely descriptive research through one quantitative questionnaire. As the study’s phases progressed, results from the previous phase informed the design of the next phase. The goal of this three-stage

---

\(^1\) Midwestern University is the pseudonym applied to my research site to maintain confidentiality of participants. All identifying information for participants was removed in reported findings.

\(^2\) In all phases, participants were visually and/or audibly presented with an informed consent form.
research process was to elaborate on student perceptions of academics during the pandemic and to vocalize each student’s experience as an individual.

This study applied mixed methods with the goal of triangulation to validate self-reported responses from the student body of Midwestern University via qualitative interviews and two separate questionnaires. All qualitative data were analyzed using qualitative content analysis in ATLAS.ti. All quantitative data were cleaned and analyzed using statistical procedures in JASP and SPSS.

Phase 1: Pilot Survey

Phase 1 consisted of a pilot survey hosted on Qualtrics, with topics regarding the participant’s perception of the American societal response to the pandemic, individual academic and social impacts of the pandemic, prevalence of clinical mental health care, and mental health resource use and needs. Using convenience sampling, the pilot survey collected 54 unique responses (45 completed, 9 partial) during May 2021. Participants were recruited through nonprobability convenience sampling, which participants were made aware of by word of mouth and social media platforms. The questionnaire included a mixture of open- and closed-ended questions in order to identify emergent trends. The pilot survey took approximately 15 minutes to complete.

Among participants, 35 (64.8%) were women, 9 (16.7%) were men, and 1 (1.9%) declined to identify their gender. Most participants (n = 39; 77.2%) identified as White, four (7.4%) identified as two or more races, and one (1.9%) identified as Hispanic/Latino. Participant ages ranged from 19 to 24 years old, and most (n = 34; 63.0%) reported being 20 or 21 years of age. Although academic year ranged from first year to graduate, most participants (n = 20; 37.0%) were second-year students during March 2020.

Phase 2: Qualitative Interviews

Phase 2 explored the emergent trends identified from pilot survey responses during May–August 2021. During Phase 2, 14 participants—11 women (78.6%) and 3 men (21.4%)—were selected for qualitative interviews using a semistructured interview guide. Interviewees were recruited through convenience sampling. No race, age, or additional demographics were recorded if participants met the criteria for eligibility. All participants were enrolled at Midwestern University before March 2020.

---

3 Triangulation allows for multiple methods of research to be utilized together to reduce possible limitations that they present individually (Dawadi, Shrestha, and Giri 2021).
4 Individual response data were analyzed only if participants met the criteria of being enrolled at Midwestern University at some level during March 2020.
2020 and at the time of the interview, and all participants had been enrolled full time at the university during March 2020.

The goal of Phase 2 was to develop a deeper understanding of student behaviors and perceptions regarding social and academic lives before and after March 2020, peer socialization patterns after March 2020, and personal responses to the pandemic such as mental health, changing norms, and societal adaptations to pandemic-induced anomie. Probing was used to further develop questions, meaning individuals could be asked to continue to elaborate on their answers if they mentioned a topic that warranted further discussion. On average, interviews lasted one hour, and all were conducted via Zoom.

Interviews were transcribed through Zoom automatic transcription and then thoroughly checked against an audio recording in order to produce the most accurate final transcript. Through open coding, emergent codes from the pilot survey were applied to these transcripts and new relevant themes were added. As a result of axial and selective coding, subsequent code groups were created, and a theoretical structure for these themes emerged. These themes informed the design of the instrument used in Phase 3.

Phase 3: Questionnaire

Phase 1 and 2 findings culminated in the development of the Phase 3 Qualtrics questionnaire. The link was distributed by departments, professors, and students through email and word of mouth. As in other phases, each participant was required to have been a part- or full-time university student during March 2020; participants had either dropped out, graduated, or continued schooling since that time. The questionnaire took approximately 17 minutes to complete.

Phase 3 research had a sample size of 253 after responses were sought from the 3,470 students eligible to participate as enrolled full time as sophomores through sixth-years at Midwestern University during May 2021. A sample size with a 95% confidence interval translates to 355 participants as calculated using the Raosoft sample size calculator. This study included 253 valid participants, creating a 94% confidence interval with 6% margin of error rather than the professional standard of 5%, thereby slightly reducing the impact of potentially statistically significant results.

A total of 183 (72.3%) women, 47 (18.6%) men, 3 (1.2%) gender-nonconforming participants, and 2 (0.8%) participants who preferred not to answer the gender

---

5 Open coding is a process in which one takes the overall text and breaks it into manageable smaller chunks of information of interest to the researcher, often by question or topic (Neuman 2020).
6 Axial coding is drawing connections between identified codes within the document (Neuman 2020).
7 Selective coding is taking multiple codes and identifying a main theme that has come forward through the research process (Neuman 2020).
question participated in this phase. The majority (79.1%) of the 200 participants identified as White, but 23 (9.1%) identified as two or more races, 7 (2.8%) identified as Asian American, 4 (1.7%) identified as Black or African American, and 1 (0.4%) identified as Hispanic/Latino. Most participants ($n = 103; 40.7\%$) were 21 years old when they completed the Phase 3 questionnaire and had been second-year or sophomore students during the March 2020 semester. All participants had the opportunity to enter their names in a gift card raffle for one of eight $25$ digital Amazon gift cards. Gift cards were funded via a grant from the Midwestern University Honors Program.

Results

This research examined how individual (micro-level) and societal (macro-level) forces interacted to shape one another. Overarching patterns related to the concepts of individual and group strain, anomie, and resilience are seen throughout the results. Emotionally, participants from Midwestern University seemed to experience intense emotional states of frustration, anger, confusion, and exhaustion due to isolation and strain on their identities and beliefs. To combat these emotional states, during Phase 2 interviews, participants modified or qualified their own actions with various reasons when discussing their adaptation and identity changes to fluctuating norms such as societal expectations, COVID-19 safety, and other emerging societal issues that were seen throughout the pandemic.\(^8\) The following sections discuss results from all phases and are grouped by emergent themes in the data with indication to the phases referenced. These themes included institutional strain, academic strain, social cohesion, mental health outcomes, societal change and anomie, and behavioral or emotional inconsistencies.

Institutional Strain: Phase 2 Data

Midwestern University chose to remove breaks (i.e., fall and spring breaks) from the Fall 2020 and Spring 2021 semesters in an attempt to reduce student travel. Instead of these breaks, the university created a “Wellness Weekend,” when professors were asked not to assign homework or tests on a chosen Friday and Monday so students could take a break. The university gave students two days of mental health programs, resources that students could take back to their rooms to complete individually, and events. Participants perceived these events as performative or lacking in authenticity and genuine action by the university toward the student body.

\(^8\) Vaccine hesitancy, politicization of the pandemic, George Floyd’s murder, Black Lives Matter, Asian American Pacific Islander hate crimes, and more are considered examples of emerging social issues that are not covered directly here.
Participants did not completely trust Midwestern University as an academic institution. Rather, participants perceived their worth from the university's perspective as quantifiable in dollars and cents—and they felt that emotional weight more than before March 2020. A quote from Jessica encapsulated this idea of economic worth over intrinsic value. Jessica, a participant enrolled in a dual-degree program with Midwestern University and Neighbor University,\(^9\) stated:

I think Midwestern University could have supplemented [no breaks] either with, you know, an actual day off like Neighbor University, [who] gave everybody two days off just randomly throughout the spring semester. Meanwhile, we [at Midwestern University] got a “Wellness Weekend.” And I still had a test that Wednesday, so I spent that weekend studying for that [class’s] test, so it’s just like they seem to be well intentioned, but it seemed more so for publicity and show for them than actually sitting there thinking of students. (emphasis added)

Sentiments expressed by Jessica and others embodied shared negative feelings of anger, frustration, and distrust of the university’s intentions and actions.

**Academic Strain: Phase 1 and 3 Data**

When asked in Phase 1 to consider their educational experience before the pandemic, participants \((n = 29; 53.7\%)\) reported a positive view of the academic system. When asked similar questions during Phase 3 about overall pre-pandemic satisfaction with their specific university, most participants reported being extremely satisfied \((n = 82; 32.4\%)\) or moderately satisfied \((n = 113; 44.7\%)\). Regarding satisfaction with the university since the pandemic had begun, participants reported that they were somewhat satisfied \((n = 76; 30\%)\), somewhat dissatisfied \((n = 58; 22\%)\), and moderately satisfied \((n = 54; 21.3\%)\). This signals a strained university-student relationship after March 2020.

**Social Cohesion: Phase 1 and 2 Data**

In response to open-ended questions in Phase 1, participants ascribed blame to citizens by using dichotomous terms such as “cared/did not care” about others’ health. In response to a question about whether they saw American society as more unified or divided post-pandemic, most participants \((n = 47; 87.0\%)\) mentioned an increased sense of division within US society. These politically charged answers and accusatory tones toward others suggest that participants may hold beliefs that could make it more difficult to unify as a group post-disaster.

\(^9\) Neighbor University and Midwestern University share a dual-degree program in which students take classes from each school to earn a degree that neither university could sustain alone.
Mental Health Discussions in Classes

Results from Phase 3 show an increase in discussion of mental health after the pandemic. Respondents were asked about mental health discussions in class across four semesters of attendance at Midwestern.\(^\text{10}\) Participants reported mental health discussion prior to Spring 2020 as occurring sometimes \((n = 123; 51.9\%)\) and never \((n = 85; 35.9\%)\), the least frequent options respondents could choose. During Spring 2020, the number lowered for sometimes \((n = 87; 36.7\%)\) but increased for about half the time \((n = 50; 21.1\%)\), most of the time \((n = 42; 117.7\%)\), and always" \((n = 24; 10.1\%)\), reducing the number for never \((n = 34; 14.3\%)\). The response for during Spring 2020 was relatable to that for Fall 2021, with a slight decrease in discussion seen for Spring 2022, but not as low as prior to Spring 2020. This suggests a large increase of discussion of mental health surrounding the outbreak of COVID-19, with a decline in Fall 2021 and Spring 2022 but still occurring more frequently than before Spring 2020. (See Tables A11 to A14.)

Mental Health Outcomes: Phase 1, 2, and 3 Data

Findings from the first and third phases of research supported the use of mental health resources by participants at Midwestern University. During Phase 1, nearly two-thirds \((n = 31; 65.9\%)\) had utilized a therapist for treatment at some point in their lives. In Phase 2, most \((n = 8; 57.1\%)\) were currently seeking or had recently sought therapy for mental health issues; however, only a few mentioned COVID-related depressive feelings, suicidal thoughts, thoughts of self-harm, or feelings of detachment from self and ascribed these feelings in whole or in part to circumstances due to the pandemic.\(^\text{11}\) In Phase 3, nearly half of participants \((n = 113; 46.9\%)\) were being treated for mental health issues in some capacity. When asked if they had ever considered seeking help from a mental health professional after March 2020, more than two-thirds of participants \((n = 125; 69.4\%)\) said yes.

A sizable number of participants \((n = 13; 24.0\%)\) in Phase 1 reported that they did not want to seek treatment, for a variety of reasons, including that their problems were not important enough, they could handle their problems on their own, or they did not have the financial or physical means to reach a therapist. Similar themes were discussed in Phase 2, including that their problems were not important enough, their

\(^{\text{10}}\) The four questions’ timeframes were before Spring 2020, during Spring 2020, during Fall 2021, and during Spring 2022, with response options of Always, Most of the time, About half the time, Sometimes, and Never.

\(^{\text{11}}\) The feelings were all discussed in the past tense and were described as times that had been difficult, mostly at the beginning of the pandemic, and no present suicidal ideations were expressed.
problems could be handled with existing coping strategies, or there was a lack of accessibility to a therapist because of financial status or geographical limitations.\textsuperscript{12}

Across all phases, participants noted accessibility barriers to university resources. In Phase 1, one-third of participants reported mental health resources on campus to be moderately accessible (\(n = 16; 29.6\%\)), and fewer reported them somewhat accessible (\(n = 11; 23.9\%\)). In Phase 2, participants reported both positive and negative interactions with university counseling staff and services; in these responses, experiences appeared to differ because of individual reasons, without an underlying factor identified as determining better or worse treatment. In Phase 3, participants reported overall campus accessibility to mental health resources as slightly (\(n = 29; 27.1\%\)) or moderately accessible (\(n = 39; 36.4\%\)) and the quality of mental health resources as average (\(n = 54; 45.4\%\)).

Across all phases of research, participants generally reported worsening mental health due to the pandemic and less ability to socialize. In a write-in option for Phase 1, some participants (\(n = 8; 14.8\%\)) cited isolation, loneliness, and uncertainty for the future, as well as general stress regarding work and schooling, as reasons for this change. In Phase 2, participants noted an increase in personal feelings of stress, anxiety, fear, isolation, emotional numbness, and inability to fully take in emotions and feelings. In Phase 3, participants reported negative emotions including restlessness (\(n = 166; 68.9\%\)), sadness (\(n = 168; 69.4\%\)), worry (\(n = 201; 83.1\%\)), fear (\(n = 142; 58.9\%\)), and confusion (\(n = 145; 59.9\%\)). Amid the negative, some participants reported positive mental health responses to the pandemic. Across all three phases, participants pitted their feelings of increased stress, isolation, and lack of socialization against the increased time to reflect and recover individually during lockdowns. In Phase 3 specifically, participants noted about the same amount of happiness (\(n = 128; 52\%\)) and contentedness (\(n = 114; 47\%\)) since the pandemic began.

\textit{Coping Mechanisms}

When students were sent home during Spring 2020, many returned to towns and cities that were not near Midwestern University. This geographic separation meant that some participants could not use university counseling because they had crossed state lines. Participants who were most affected were those in Southern and far Northern Midwest states, who experienced a technological, physical, and economic lack of accessibility to many university services. In Phase 2, those who could not attend therapy because of financial, geographical, or related accessibility barriers coped by using exercise and resources they already had at their disposal, such as puzzles, games, painting, or video games. Importantly, those in less-affluent and

\textsuperscript{12} Limitations included those associated with rural areas’ limited Internet access and long travel distances to access resources.
remote areas also mentioned barriers to Internet accessibility, which increased isolation because of a lack of socialization with university friends.

When students returned to campus in Fall 2020 and Spring 2021, participants from Phase 2 interviews alluded to coping with problems by registering an emotional support animal (n = 1; 7.14%), using outside counseling services (n = 5; 35.7%), participating in hobbies (n = 5; 35.7%), exercising (n = 1; 7.1%), communicating with friends (n = 14; 100%), communicating with family (n = 8; 57.1%), consuming alcohol (n = 2; 14.2%), partying (n = 2; 14.2%), and “ignoring the problem” (n = 2; 14.2%). In Phase 3, participants reported using the following mental health resources that the university provided or informed them about: individual counseling (n = 94; 79.0%), a university-subsidized app for mental health (n = 39; 32.8%), faith-based programming (n = 15; 12.6%), university group counseling (n = 13; 10.9%), university-recommended outside providers (n = 10; 8.4%), mental health hotlines (n = 4; 3.4%), mental health awareness programming (n = 2; 1.7%), and “Other” (n = 2; 1.7%).

**Stress, Anxiety, and Control**

In Phase 1, participants reported varying levels of change in their stress levels since the pandemic began: a great or extreme amount (n = 33; 61.1%), a moderate amount (n = 11; 20.4%), and a small amount (n = 1; 1.8%). Also in Phase 1, participants reported higher anxiety (n = 39; 72.2%), a higher perceived lack of control over their lives (n = 36; 66.7%), and about the same (n = 13; 24%) or lower (n = 27; 50%) motivation to complete tasks compared to before March 2020. In Phase 3, anxiety was not directly assessed, but participants responded that since the beginning of the pandemic, they had experienced negative changes in control over their own lives (n = 176; 73%; p < .05). During Phase 3, 190 participants (81%) felt that stress had increased since the pandemic began in March 2020.

**Isolation**

Phase 2 findings suggested the important role of isolation in mental health. For example, Sherry, one of the Phase 2 interviewees, described feeling isolated when students were sent home from school during Spring 2020. She said “and so it was like we could either be on Zoom or we could be alone, and I think everybody just chose to be alone” (emphasis added). The role of isolation varied among the 14 interviewees, but when mentioned, isolation was always tied to the pandemic. One (7.14%) had been ostracized from their friends because of their COVID-19 policy beliefs, some (n = 4; 28.57%) felt trapped in their on- or off-campus housing, some (n = 4; 28.57%) felt afraid to socialize, and some (n = 1; 7.14%) felt targeted for being in Greek Life. These

---

13 Midwestern University Greek Life students were hosting parties in off-campus housing or Greek housing, and this was associated with a spike in campus COVID cases. The negative stigma then grew
struggles related to changing social norms and a clash of individual beliefs and identities. Results from Phase 2 suggested that isolation was seemingly tied to other negative emotions, such as fear and anxiety, loneliness, helplessness, and frustration.

**Frustration**

Some Phase 2 participants \((n = 13; 24.1\%)\) felt an increase in personal frustrations during the pandemic, such as feeling overwhelmed because of being asked to change a lot of behavior or not being valued enough by others (e.g., their university, friend groups, or communities) after returning to campus. Participants provided a variety of reasons for this frustration, including challenging their political affiliations, limiting personal freedoms, sharing/not sharing personal beliefs among peers, and academic struggles. Such negative emotions were echoed in Phase 3, when participants were asked about experiencing more or less anger since the pandemic began. To this question, half of participants \((n = 121; 50.0\%)\) reported feeling more anger.

**Societal Change and Anomie: Phase 1 and 2 Data**

Within Phase 1, participants used “pre-” and “post-COVID” language, which suggested a change in society norms and values. When asked directly, 49 of 54 (90.7\%) participants said at least some aspect of society had changed because of the pandemic, tentatively confirming that there was a change or restructuring of social values. Most participants in Phase 1 \((n = 43; 79.6\%)\) reported that an aspect of their life had changed as part of the pandemic, and only four participants (7.4\%) said an aspect of their life had not changed. All Phase 2 participants reported some impact within their academic, social, and familial circles, and nearly all expressed increased COVID-related anxiety \((n = 13; 92.9\%)\), fear \((n = 13; 92.9\%)\), and isolation \((n = 14; 100.0\%)\). Additionally, in Phase 2, most participants \((n = 10; 71.4\%)\) expressed that there had been times they had felt an increase in heavy emotional situations or pandemic-related strain.\(^{14}\) These responses suggest an increase in the feelings of individual societal strain related to anomie.

A Phase 2 participant named Rain demonstrated an interesting paring of viewpoints about distancing and social norms from her experience as being Latina. Rain highlighted an aspect of American culture that made the adjustment to new social distancing norms easier because of Americans being more distant socially

\(^{14}\) This code is used when the participant described (or directly said) something large (emotionally or mentally) that was putting a lot of weight on their psyche.
before the pandemic. She compares this with Mexican culture where it is more normal to sit or stand much closer to another person. Since the pandemic, she had noticed this specific change within society, pointing out a now unified norm of distancing within American culture. Phil, another Phase 2 participant, noted a social norm change that was influenced by the university community:

It was kind of interesting I think. I have a couple of friends that go to Northwestern in Chicago, and then I have a friend that goes to university in Georgia. I have a couple of friends that go to Wake Forest in North Carolina. And so, kind of taking those four schools—those three schools that I mentioned and Midwestern University—comparing the different experiences, it truly felt like they were all four dealing with a different virus.

Phil then explained how between these four schools, in-person and online classes differed, with some students very restricted to their dorms, and some were strict regarding rules for masks or being outside for activities. Although Phil noticed these differences, however, he did not attribute the differences to any specific social or cultural differences among educational institutions.

Behavioral and Emotional Inconsistency: Phase 1, 2, and 3 Data

Behavioral Inconsistency

Within Phase 2, participants spoke to a sense of obligation and needing to “do your part” to protect others by adhering to behavioral changes to norms (e.g., mask wearing and social distancing). Some participants \((n = 5; 9.2\%)\) discussed being uncomfortable and angry when their peers or roommates were not being careful or respecting others’ health by not wearing masks or socializing with certain friends. Participants spoke about giving up normal activities for the safety of others, such as seeing people in person, seeing certain friends, or going to stores, and these acts weighed heavily on them. Across all phases, participants discussed concerns over others’ or their own health as potentially influencing their behavior to attend or avoid in-person gatherings. Overall, this speaks to an influence of the health of others, rather than the health of oneself, as a strong indicator of personal behavioral motivations. This centers around a theme of the emotional guilt and weight that participants put on themselves or took in from social situations regarding the uncertain social norms and personal belief of how others should act. These feelings and beliefs could have led to further division rather than cohesion once students were back on campus in Fall 2021, possibly reducing student resilience. These findings could support a lack of social cohesion among students in a post-pandemic environment.
Emotional Inconsistency during the Pandemic

In Phase 3 findings, students reported feeling overwhelmed by their academic expectations since 2020 (n = 226; 89.3%). According to participants, being overwhelmed influenced their mental health (n = 215; 96.8%) a great deal (n = 74; 34.3%) or a lot (n = 76; 35.3%) in a negative manner (n = 161; 97.6%). Specifically, participants felt overwhelmed by in-class expectations (n = 170; 70.2%), which affected their mental health (n = 146; 90.1%) a moderate amount (n = 56; 38.9%) or a lot (n = 39; 27.1%) or a great deal (n = 39; 27.1%) in a negative manner (n = 126; 98.4%). Participants also reported feeling overwhelmed by professors’ expectations (n = 178; 73.9%), which influenced their mental health (n = 146; 86.4%) a moderate amount (n = 50; 34.3%) or a great deal (n = 42; 28.8%) in a negative manner (n = 117; 97.5%). Finally, participants reported academic pressure had influenced their mental health (n = 117; 88.0%) a lot (n = 48; 41.4%) or a great deal (n = 40; 34.5%). However, participants were relatively divided on whether they felt increased academic pressure (n = 113; 45.4%) or a similar sense of pressure (n = 109; 43.8%) to succeed academically compared to before the pandemic.

In Phase 3, a chi-squared test was run between the options “negative changes in their lives since the pandemic began” and five individual responses of control over their own lives, emotional capacity, stress, perspective on life, and perception of self, which generated significant and nonsignificant results. These results show that since the start of the pandemic, participants had experienced negative changes in control over their own lives (n = 176; 73%; p < .05), emotional capacity (n = 136; 56%; p < .05), stress (n = 190; 81%; p < .05), perspective on life (n = 111; 46%), and perception of self (n = 105; 44.5%).

Participants reported feeling more of the following negative emotions since the pandemic began: restlessness (n = 166; 68.9%), sadness (n = 168; 69.4%), worry (n = 201; 83.1%), fear (n = 142; 58.9%), and confusion (n = 145; 59.9%). After March 2020, participants reported, they felt overwhelmed because of professor expectations, academic expectations, and in-class expectations, yet participants also reported that since the pandemic started, their personal academic standards were “about the same” (n = 116; 46.8%). These findings suggest a potential increase in strain resulting from increased pressure within academics and without a corresponding decrease in personal expectations regarding academic success; however, no open-ended follow-up question was given to participants.

Discussion

In this section, I discuss the implications of this study’s findings as they relate to anomie, social strain, mental health, and student resources. Results from these three phases of research align with previous post-disaster response literature in

Each phase of this study examined the topics of student mental health, accessibility of resources, and adaptative responses during and after the pandemic. The theory of anomie describes the social response to an absence of norms that regulate behaviors. The COVID-19 pandemic brought about a sense of anomie on university campuses because of shifting social norms, a lack of access to resources, and increased emotional strain. Given the many changes that students experienced during the pandemic, findings suggest that, because of prolonged fluctuation of social norms, participants struggled to reconcile their social identities as members of a group (i.e., Midwestern University students) and as individuals.

Anomie was clearly present when participants discussed a lack of understanding of what to do in social settings because of differences in expectations before and after the pandemic. Shifting norms, or a lack of norms, produced strain in students’ academic and individual lives. Friendships were tested by pandemic-related behaviors, such as wearing masks, physically distancing, and so on. This increased awareness of norms brought forth conflict and confusion, thereby leading to strain for the individual.

Participants in the current study expressed patterns of anxiety, isolation, stress, frustration, fatigue, and confusion due to the pandemic’s influence on their lives, suggesting that these may be persistent feelings associated with the pandemic. In addition, participants repeatedly highlighted their concerns about the need for social support networks and resources. Work from the American College Health Association (ACHA, 2019); Becker (2020); Fruehwirth, Biswas, and Perreira (2021); and Usher et al. (2020) produced results similar to the current study’s findings regarding the emotional responses of participants (relating to the pandemic).

In the first two phases of this study, participants reported using adaptative coping strategies to manage their feelings, with partial to full success, which mirrored the resilience tactics used by other communities after past disasters (Almazan et al. 2018; Shailaja et al. 2020; Shing, Jaywickreme, and Waugh 2016). Specifically, these students’ mental health and emotional needs aligned with findings that more access to health-related resources decreased negative mental health needs. Participants also felt they were still not given proper support in the academic and individual spheres, as demonstrated by references to increased academic expectations and strain, inadequate support from Midwestern University for burnout and fatigue, and heightened awareness of students’ economic value to the university. As participants reported less perceived support by the administration (i.e., seeing themselves as economic pawns), they may have struggled to reconcile their social identities as Midwestern University students with their individual identities as people, increasing
the isolation and confusion they felt on top of academic pressures. This may have reduced students’ resilience when they returned to campus for Fall 2020, because of a less cohesive and shared group identity of students at Midwestern University.

Mental health coping mechanisms were often used to reduce cognitive dissonance that resulted from changing norms. Although results of this study cannot be generalized to the macro level, findings suggest that institutions of higher education, specifically Midwestern University, would have benefitted from additional resources to support student mental health. Trends surrounding mental health and accessibility to resources may correspond with resiliency, or a lack thereof, during the pandemic. The perceived lack of access to high-quality mental health services discussed in each phase of research highlights the need for more affordable mental health services in higher education. With funding to increase accessibility of mental health services, on campuses or in communities, the use of these resources would likely increase. In turn, stigma associated with mental health resources may decrease.

In interviews throughout Phase 2, participants suggested that remote areas lacked resources, which negatively affected individuals’ abilities to cope and be resilient. Policies that allowed universities to lower costs to access all student services could help improve student mental health following a disaster. More direct access when off campus could help students’ well-being and improve their social support systems. If this disparity in access is addressed, students within these communities may experience lower levels of isolation, more equitable access to information necessary for scholarly work, and a reduction in stress-related health problems. Provision of equity in technological accessibility could be an important step in reducing feelings of isolation, stigma, and normlessness in future pandemics.

Because participants perceived a heavier academic load, upon returning to campus, the implementation of a “Wellness Weekend” in place of a sustained break appeared to increase strain. Some specifically mentioned that a traditional and longer break could have reduced student stress. Participants suggested that reducing the cognitive load for one class by watching a documentary or canceling class would have been helpful. Next, participants referenced a need to reward students for good behavior and academic performance by providing positive reinforcement rather than performative care, thereby potentially improving self-esteem. The final suggestion offered by participants was simply asking how a student was doing.

The COVID-19 pandemic exacerbated challenges already present within higher education: inequitable access to resources, the business model of higher education, and worsening mental health. Given that educational institutions are a

---

15 During Phase 2 interviews, students mentioned feeling strong support or lack of support from professors, which caused either better or worse coping for the student. This seemed professor-specific or student-specific rather than relating to any overall trends.
microcosm of broader society, it is not surprising that university students struggled with resiliency during the pandemic. If university campuses, and society more broadly, continue to erode social cohesion, anomie will continue to be a rational response to these social strains.

Conclusion

Participants within this study demonstrated a need for community and university support that was partially met during the pandemic. Both structural and individual factors coincided to decrease feelings of well-being among students. Research has consistently shown that, to promote resilience following a disaster, a community’s basic needs, such as food and water, must be met; after those needs are met, positive contextual coping (i.e., encouraging positive emotions in order to mitigate stress) and other strategies should be provided to assist communities (Shing, Jaywickreme, and Waugh 2016). Such positive emotions may stem from shared religious beliefs, shared identities, feelings of trust and unity, or a shared group identity (Almazan et al. 2018; Drury et al. 2015; Maki et al. 2018; Prosser et al. 2020; Shing, Jaywickreme, and Waugh 2016).

If institutions of higher education can facilitate an environment conducive to positive contextual coping, students may feel more supported and respond more positively to challenges associated with a pandemic. Ideas for facilitating such an environment include freezing tuition costs, increasing financial aid, increasing outreach by campus religious and non-faith-based groups, and improving communications that display a sense of unity.
References


Esterwood, Emily, and Sy Atezaz Saeed. 2020. “Past Epidemics, Natural Disasters, COVID19, and Mental Health: Learning from History as We Deal with the


## Appendix

### Multiphase Statistics Tables

<table>
<thead>
<tr>
<th>Table A1. Participants Who Indicated Anxiety, Fear, Isolation, or Strain during Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Maggie</td>
</tr>
<tr>
<td>Sherry</td>
</tr>
<tr>
<td>Jessica</td>
</tr>
<tr>
<td>Rachel</td>
</tr>
<tr>
<td>Phil</td>
</tr>
<tr>
<td>Catherine</td>
</tr>
<tr>
<td>Stephanie</td>
</tr>
<tr>
<td>William</td>
</tr>
<tr>
<td>Rain</td>
</tr>
<tr>
<td>Catrina</td>
</tr>
<tr>
<td>Toby</td>
</tr>
<tr>
<td>Marcia</td>
</tr>
<tr>
<td>Emmy</td>
</tr>
<tr>
<td>Emily</td>
</tr>
<tr>
<td><strong>Total (n)</strong></td>
</tr>
</tbody>
</table>

92.9% 92.9% 100.0% 71.4%

**Total (N) =14**
Table A2. Responses to “Throughout the pandemic, to what extent did concerns about your health influence your behaviors to attend or avoid in-person gatherings?” (by academic level in March 2020)

<table>
<thead>
<tr>
<th>Responses</th>
<th>First Year</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>A great deal</td>
<td>13</td>
<td>17.8</td>
<td>34</td>
<td>30.9</td>
<td>9</td>
<td>37.5</td>
<td>2</td>
</tr>
<tr>
<td>A lot</td>
<td>14</td>
<td>19.2</td>
<td>14</td>
<td>12.7</td>
<td>3</td>
<td>12.5</td>
<td>4</td>
</tr>
<tr>
<td>A moderate amount</td>
<td>19</td>
<td>26.0</td>
<td>28</td>
<td>25.5</td>
<td>9</td>
<td>37.5</td>
<td>2</td>
</tr>
<tr>
<td>A little</td>
<td>23</td>
<td>31.5</td>
<td>22</td>
<td>20.0</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
</tr>
<tr>
<td>None at all</td>
<td>4</td>
<td>5.5</td>
<td>12</td>
<td>10.9</td>
<td>1</td>
<td>4.2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>73</td>
<td>100.0</td>
<td>110</td>
<td>100.0</td>
<td>24</td>
<td>100.0</td>
<td>10</td>
</tr>
</tbody>
</table>
Table A3. Chi-Squared Test for Crosstabulation: Concerns of Health of Self

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic significance (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson chi-square</td>
<td>33.467</td>
<td>20</td>
<td>.030</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>30.412</td>
<td>20</td>
<td>.063</td>
</tr>
<tr>
<td>Linear-by-linear association</td>
<td>.940</td>
<td>1</td>
<td>.332</td>
</tr>
<tr>
<td>No. valid cases</td>
<td>220</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A4. Responses to “Throughout the pandemic, to what extent did concerns about your health influence your behaviors to attend or avoid in-person gatherings?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>A great deal</td>
<td>60</td>
<td>23.7</td>
<td>27.0</td>
</tr>
<tr>
<td>A lot</td>
<td>35</td>
<td>13.8</td>
<td>15.8</td>
</tr>
<tr>
<td>A moderate amount</td>
<td>59</td>
<td>23.3</td>
<td>26.6</td>
</tr>
<tr>
<td>A little</td>
<td>48</td>
<td>19.0</td>
<td>21.6</td>
</tr>
<tr>
<td>None at all</td>
<td>20</td>
<td>7.9</td>
<td>9.0</td>
</tr>
<tr>
<td>Missing</td>
<td>31</td>
<td>12.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table A5. Responses to “Throughout the pandemic, to what extent did concerns about others’ health influence your behaviors to attend or avoid in-person gatherings?” (by academic level in March 2020)

<table>
<thead>
<tr>
<th>Response</th>
<th>First Year</th>
<th>Second Year</th>
<th>Third Year</th>
<th>Fourth Year</th>
<th>Fifth Year</th>
<th>Sixth Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>A great deal</td>
<td>3</td>
<td>45.2</td>
<td>48</td>
<td>43.6</td>
<td>14</td>
<td>58.3</td>
</tr>
<tr>
<td>A lot</td>
<td>16</td>
<td>21.9</td>
<td>30</td>
<td>27.2</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>A moderate amount</td>
<td>16</td>
<td>21.9</td>
<td>17</td>
<td>15.4</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>A little</td>
<td>5</td>
<td>6.8</td>
<td>12</td>
<td>10.9</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>None at all</td>
<td>5</td>
<td>6.8</td>
<td>3</td>
<td>2.7</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>100.0</td>
<td>110</td>
<td>100.0</td>
<td>24</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table A6. Chi-Squared Test for Crosstabulation: Concerns of Health of Others

<table>
<thead>
<tr>
<th>Response</th>
<th>Value</th>
<th>df</th>
<th>Asymptotic significance (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson chi-square</td>
<td>14.783</td>
<td>20</td>
<td>.789</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>14.528</td>
<td>20</td>
<td>.803</td>
</tr>
<tr>
<td>Linear-by-linear association</td>
<td>.111</td>
<td>1</td>
<td>.739</td>
</tr>
<tr>
<td>No. valid cases</td>
<td>220</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A7. Response to “Throughout the pandemic, to what extent did concerns about other’s health influence your behaviors to attend or avoid in-person gatherings?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>A great deal</td>
<td>103</td>
<td>40.7</td>
<td>46.4</td>
</tr>
<tr>
<td>A lot</td>
<td>54</td>
<td>21.3</td>
<td>24.3</td>
</tr>
<tr>
<td>A moderate amount</td>
<td>38</td>
<td>15.0</td>
<td>17.1</td>
</tr>
<tr>
<td>A little</td>
<td>18</td>
<td>7</td>
<td>8.1</td>
</tr>
<tr>
<td>None at all</td>
<td>9</td>
<td>3.6</td>
<td>4.1</td>
</tr>
<tr>
<td>Missing</td>
<td>31</td>
<td>12.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table A8. Responses to “During the Fall 2020–Spring 2021 semester please describe your attendance to the following in-person events: Gatherings with university organizations.”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased a great deal</td>
<td>8</td>
<td>3.2</td>
<td>7.1</td>
</tr>
<tr>
<td>Somewhat increased</td>
<td>18</td>
<td>7.1</td>
<td>16.1</td>
</tr>
<tr>
<td>Neither increased nor decreased</td>
<td>30</td>
<td>11.9</td>
<td>26.8</td>
</tr>
<tr>
<td>Somewhat decreased</td>
<td>56</td>
<td>22.1</td>
<td>50.0</td>
</tr>
<tr>
<td>Seen but not selected</td>
<td>18</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>123</td>
<td>48.6</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table A9. Responses to “During the Fall 2020–Spring 2021 semester please describe your attendance to the following in-person events: Gatherings with friends.”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased a great deal</td>
<td>18</td>
<td>7.1</td>
<td>11.0</td>
</tr>
<tr>
<td>Somewhat increased</td>
<td>28</td>
<td>11.1</td>
<td>17.1</td>
</tr>
<tr>
<td>Neither increased nor decreased</td>
<td>50</td>
<td>19.8</td>
<td>30.5</td>
</tr>
<tr>
<td>Somewhat decreased</td>
<td>68</td>
<td>26.9</td>
<td>41.5</td>
</tr>
<tr>
<td>Seen but not selected</td>
<td>16</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>73</td>
<td>28.9</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table A10. Responses to “During the Fall 2020–Spring 2021 semester please describe your attendance to the following in-person events: Gatherings with family.”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased a great deal</td>
<td>17</td>
<td>6.7</td>
<td>10.4</td>
</tr>
<tr>
<td>Somewhat increased</td>
<td>21</td>
<td>8.3</td>
<td>12.8</td>
</tr>
<tr>
<td>Neither increased nor decreased</td>
<td>48</td>
<td>19.0</td>
<td>29.3</td>
</tr>
<tr>
<td>Somewhat decreased</td>
<td>78</td>
<td>30.8</td>
<td>47.6</td>
</tr>
<tr>
<td>Seen but not selected</td>
<td>17</td>
<td>6.7</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>72</td>
<td>28.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
Table A11. Responses to “How often was student mental health discussed in your classes in the following academic time spans: Before Spring 2020?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>3</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>Most of the time</td>
<td>9</td>
<td>3.6</td>
<td>3.8</td>
</tr>
<tr>
<td>About half the time</td>
<td>17</td>
<td>6.7</td>
<td>7.2</td>
</tr>
<tr>
<td>Sometimes</td>
<td>123</td>
<td>48.6</td>
<td>51.9</td>
</tr>
<tr>
<td>Never</td>
<td>85</td>
<td>33.6</td>
<td>35.9</td>
</tr>
<tr>
<td>Missing</td>
<td>16</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table A12. Responses to “How often was student mental health discussed in your classes in the following academic time spans: During Spring 2020?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>24</td>
<td>9.5</td>
<td>10.1</td>
</tr>
<tr>
<td>Most of the time</td>
<td>42</td>
<td>16.6</td>
<td>17.7</td>
</tr>
<tr>
<td>About half the time</td>
<td>50</td>
<td>19.8</td>
<td>21.1</td>
</tr>
<tr>
<td>Sometimes</td>
<td>87</td>
<td>34.4</td>
<td>36.7</td>
</tr>
<tr>
<td>Never</td>
<td>34</td>
<td>13.4</td>
<td>14.3</td>
</tr>
<tr>
<td>Missing</td>
<td>16</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table A13. Responses to “How often was student mental health discussed in your classes in the following academic time spans: During Fall 2021?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>21</td>
<td>8.3</td>
<td>8.9</td>
</tr>
<tr>
<td>Most of the time</td>
<td>51</td>
<td>20.2</td>
<td>21.5</td>
</tr>
<tr>
<td>About half the time</td>
<td>57</td>
<td>22.5</td>
<td>24.1</td>
</tr>
<tr>
<td>Sometimes</td>
<td>88</td>
<td>34.8</td>
<td>37.1</td>
</tr>
<tr>
<td>Never</td>
<td>20</td>
<td>7.9</td>
<td>8.4</td>
</tr>
<tr>
<td>Missing</td>
<td>16</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
Table A14. Responses to “How often was student mental health discussed in your classes in the following academic time spans: During Spring 2022?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>12</td>
<td>4.7</td>
<td>7.4</td>
</tr>
<tr>
<td>Most of the time</td>
<td>20</td>
<td>7.9</td>
<td>12.3</td>
</tr>
<tr>
<td>About half the time</td>
<td>29</td>
<td>11.5</td>
<td>17.8</td>
</tr>
<tr>
<td>Sometimes</td>
<td>76</td>
<td>30.0</td>
<td>46.6</td>
</tr>
<tr>
<td>Never</td>
<td>26</td>
<td>10.3</td>
<td>16.0</td>
</tr>
<tr>
<td>Missing</td>
<td>90</td>
<td>35.6</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>253</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table A15. Responses to “Since the pandemic began have you felt a positive or negative change in any of the following: Control over own life?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive change</td>
<td>21</td>
<td>8.3</td>
<td>8.7</td>
</tr>
<tr>
<td>No change</td>
<td>44</td>
<td>17.4</td>
<td>18.3</td>
</tr>
<tr>
<td>Negative change</td>
<td>176</td>
<td>69.6</td>
<td>73.0</td>
</tr>
<tr>
<td>Seen but not selected</td>
<td>1</td>
<td>.4</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>11</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>253</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table A16. Responses to “Since the pandemic began have you felt a positive or negative change in any of the following: emotional capacity?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive change</td>
<td>35</td>
<td>13.8</td>
<td>14.5</td>
</tr>
<tr>
<td>No change</td>
<td>70</td>
<td>27.7</td>
<td>29.0</td>
</tr>
<tr>
<td>Negative change</td>
<td>136</td>
<td>53.8</td>
<td>56.4</td>
</tr>
<tr>
<td>Seen but not selected</td>
<td>1</td>
<td>.4</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>11</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>253</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table A17. Chi-Squared Test for Crosstabulation: Perspective on Life

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic significance (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson chi-square</td>
<td>12.940</td>
<td>10</td>
<td>.227</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>15.174</td>
<td>10</td>
<td>.126</td>
</tr>
<tr>
<td>Linear-by-linear association</td>
<td>3.967</td>
<td>1</td>
<td>.046</td>
</tr>
</tbody>
</table>

No. valid cases 231
Table A18. Responses to “Since the pandemic began have you felt a positive or negative change in any of the following: Perspective on life?”

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive change</td>
<td>68</td>
<td>26.9</td>
<td>28.6</td>
</tr>
<tr>
<td>No change</td>
<td>59</td>
<td>23.3</td>
<td>24.8</td>
</tr>
<tr>
<td>Negative change</td>
<td>111</td>
<td>43.9</td>
<td>46.6</td>
</tr>
<tr>
<td>Seen but not selected</td>
<td>4</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>11</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table A19. Chi-Squared Test for Crosstabulation: Perception of Self

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic significance (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson chi-square</td>
<td>12.472</td>
<td>10</td>
<td>.255</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>15.330</td>
<td>10</td>
<td>.120</td>
</tr>
<tr>
<td>Linear-by-linear association</td>
<td>2.971</td>
<td>1</td>
<td>.085</td>
</tr>
<tr>
<td>No. valid cases</td>
<td>229</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A20. Responses to “Since the pandemic began have you felt a positive or negative change in any of the following: Perception of self?”

<table>
<thead>
<tr>
<th>Responses</th>
<th>Frequency (N)</th>
<th>Percent</th>
<th>Valid percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive change</td>
<td>43</td>
<td>17.0</td>
<td>18.2</td>
</tr>
<tr>
<td>No change</td>
<td>88</td>
<td>34.8</td>
<td>37.3</td>
</tr>
<tr>
<td>Negative change</td>
<td>105</td>
<td>41.5</td>
<td>44.5</td>
</tr>
<tr>
<td>Seen but not selected</td>
<td>5</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>12</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
Abstract

Research focused on the intersection of social media messages and perceptions of restorative justice has the potential to provide insight on criminal justice reform. In light of national conversations about mass incarceration, restorative justice has become a topic of conversation on social media, particularly among activists. This research utilized qualitative interviews (n = 20) to ascertain Gen Z college students’ familiarity with, understanding of, and interpretation of messages about restorative justice as portrayed in social media posts that they have been exposed to during the course of their everyday lives. The study found that most participants learned about restorative justice in the classroom and generally had positive perceptions of restorative justice, and that many believed the general public likely views restorative justice as “soft on crime.” Additionally, the severity of the crime influenced how accepting participants were of restorative justice. These findings have the potential to contribute to the literature on restorative justice, as well as to current advocacy practices as they are perceived by young adults.

Introduction

The United States has the highest incarceration rate in the world, especially among developed countries, despite using some of the harshest punishment methods and policing practices (Frost, 2006; New York Times, 2021; Pew Research Center, 2018). Pew Research Center data reveal that the United States has an incarceration rate of over 600 people per 100,000 people, significantly more than any other country (Pew Research Center, 2018). Restorative justice is a term that entered the lexicon in recent years with the media highlighting the ineffectiveness and inhumanity of the current criminal justice system. Restorative justice serves as an alternative to traditional forms of punishment such as prison sentences (Doolin 2007; Morris & Maxwell, 2001). Restorative justice aims to fundamentally change the inappropriate behavior of criminal offenders by addressing the underlying factors and conditions that might drive an individual to delinquency (Clear & Karp, 2002; Morris & Maxwell, 2001; Wright, 1996).
Restorative justice also aims to help reintegrate offenders back into society, help them take responsibility for their actions, and involve the community and victims in the process to rebuild trust and achieve closure for victims (Clear & Karp, 2002; Doolin, 2007; Morris & Maxwell, 2001; New York Times, 2021). One of the early advocates for restorative justice, Tony Marshall, defined restorative justice as the process of solving existing problems and issues by involving both parties and collectively deciding how to move forward while also dealing with the aftermath of the incident (Dignan, 2005; Marshall, 1999). Rather than focusing on punishment, restorative justice emphasizes the importance of rehabilitation and repairing harms. The future of the criminal justice system may be intertwined with restorative justice practices such as victim-offender reconciliation and conferencing, group therapy, and community involvement as a rehabilitative approach proves its effectiveness.

The United States currently has a punitive, or punishment-based, criminal justice system. The idea behind this approach is to deter crime and punish people for breaking the laws set in place (Dolovich & Natapoff, 2017; Frost, 2006). Despite the historically punitive-based approach to crime, the United States has seen its criminal justice system becoming increasingly punitive-based over recent decades. The increasingly punishment-based system has likely contributed to the issue of mass incarceration that exists within American society (Dolovich & Natapoff, 2017). The punitive criminal justice system has become stricter in terms of sentencing minimums and percentages that must be served, the use of capital punishment, and the ways in which the system punishes people (Dolovich & Natapoff, 2017). For example, the United States uses solitary confinement in prisons to punish offenders; however, studies have shown that solitary confinement is extremely detrimental to an individual’s mental health and often sets offenders back in terms of their criminality and behavior (Casella et al., 2016).

Instead of measuring success within the criminal justice system based on how released offenders function in their communities or on decreased recidivism rates, the punitive system emphasizes prosecution and conviction as measures of success (Dolovich & Natapoff, 2017). The current punitive criminal justice system also fails to address the root causes of crime and how society affects individuals negatively. Restorative justice offers a beacon of hope to offenders stuck in the punitive system and offers offenders a chance to heal and create meaningful personal change. Restorative justice takes a rehabilitative approach to crime and criminal offenders. Rather than putting people in prison and releasing them without providing them with the crucial tools required to survive after their release, restorative justice aims to equip offenders with the resources to thrive. Instead of expecting behavioral and cognitive changes to occur as a result of punishment, restorative justice practices help create meaningful change and treat offenders with humanity and dignity (Judah & Bryant, 2004).
Awareness of restorative justice appears to be increasing, although it is not a widely used practice in society. Knowing the perceptions and attitudes that the public has toward restorative justice is important because as a democratic society, citizens have the power to influence the laws and policies that affect them (New York Times, 2021). One possible way to examine the public’s perception and opinions regarding a wide range of social issues, including restorative justice, is through social media. Social media has shown its usefulness as a news source, which many organizations are capitalizing on. Studies have also found that the public tends to have positive attitudes toward certain businesses or brands after interacting with those organizations through social media (Funk, 2013). For example, one study examined how police departments utilize social media as a news outlet to reach their target audience. Cheng (2021) discovered that social media reduces reliance on mass media and also allows police departments to spin stories in their favor. Through social media, the New York Police Department was able to effectively defuse the public’s negative perceptions of contested situations, including a 2018 police shooting (Cheng, 2021). Other studies have examined how media allows people to engage in political activism without physical presence, simplifies communication and coordination, and may help amplify marginalized voices (Schroeder, 2018). It is crucial to understand the real-world implications of social media for people, as social media is often used to influence and persuade audiences. Moreover, it is important to recognize how social media may be used to portray restorative justice currently, and also how it could be used to increase awareness of restorative justice in the future.

Generation Z (hereafter referred to as Gen Z)\(^1\) is an important demographic to study when investigating the influence of social media on public perception and attitudes, because Gen Z uses social media at an extremely high rate compared to other generations of users (Duffett, 2017). Research suggests that Americans under the age of 30 use platforms such as Instagram, TikTok, and Snapchat more than other age groups do and also use social media in general more than older Americans do (Auxier & Anderson, 2021). Many businesses and organizations are targeting their virtual marketing on this age group through social media because of these high rates of usage among Gen Z individuals (Duffett, 2017; Parker & Igielnik, 2020). Social media is a powerful tool for social justice advocates and organizations, and Gen Z is often involved in the creation and engagement of content. Importantly, studies suggest that younger people use social media to engage in political participation and advocacy more than older people (Schroeder, 2018; Tyson et al. 2021). Similar to social media, television and news outlets have a large impact on the public’s perception of crime and safety in their area (Eschholz et al., 2002; Rosenberger & Callanan, 2011; Weitzer & Kubrin, 2004).

---

\(^1\) Gen Z includes individuals born after 1996 (Parker & Igielnik, 2020).
The purpose of this case study is to understand through qualitative interviews how Gen Z students at a medium-sized Midwestern university interpret social media messages portraying restorative justice. Restorative justice is defined as the practices that aim to restore community and victims’ trust in the offender, understand the motivating factors that cause people to commit crimes, help the victim get closure and answer questions about their crime, engage the offender, encourage the offender to take responsibility for their actions, and help reintegrate the offender back into their community and society (Clear & Karp, 2002; Doolin, 2007; Morris & Maxwell, 2001; New York Times, 2021; Wright, 1996). Messages of restorative justice are defined as social media posts that discuss alternatives to traditional forms of coercive punishment for offenders or, more broadly, critique the criminal justice system in its current form. Such posts may be created by individuals or activist organizations on any social media platform, including Instagram, Twitter, Facebook, and TikTok.

Importantly, one poll found that 64% of respondents support restorative justice, and the percentage increased to 70% when easy-to-understand language was used (American Friends Service Committee, 2021). Despite generally high public support, restorative justice has not been widely implemented in the American criminal justice system. Confusion surrounding what restorative justice is and its practices likely prevent an increase in public support. This study aims to examine the complexities regarding the public’s perception of restorative justice, how restorative justice has been portrayed in the media, and what factors into opinions of restorative justice.

Methods

I conducted 26 total semistructured interviews between November 2021 and February 2022. Six of those interviews were excluded from analysis because of the participants’ unfamiliarity with restorative or community based justice. Some participants confused it with related terms (e.g., social justice, Black Lives Matter movement, racial equality movements), had looked up the term before the interview, or simply were not familiar with the term. When a participant did not know what restorative justice was or had never heard of it before, I ended the interview and deleted any collected data (e.g., interview recording). If a participant had a vague understanding, such as having heard of the term but being unable to fully define it, I proceeded with the interview and asked questions that were relevant to the participant’s knowledge and made a note on the transcript document. Occasionally, I did not realize a participant was confusing restorative justice with a different social justice movement or term until the interview was partially completed; in these cases, I continued with the interview and made a note at the end of the transcript.
Coding

I began the open coding process by reading through each transcript in its entirety before beginning to code. Through open coding, I began coding by examining the major themes I found throughout the transcript, in order to narrow the codes into specific categories or groups. After transcription, the interviews were inductively coded to identify valuable themes and patterns as they emerged from data (Thomas, 2006). This emergent coding approach allowed me to create new codes as the themes and patterns emerged from data, rather than creating codes prior to beginning the coding process (Neuman, 2011; Thomas, 2006). A few of the most common themes throughout the transcripts included the effectiveness of restorative justice, the source or context in which the participant had become familiar with restorative justice, and possible demographic factors that might influence the public’s perception of restorative justice.

Sample

The sample consisted of 20 participants from a predominantly White, medium-sized Midwestern university, consisting of 19 females and 1 male. Each participant was given the option to choose a pseudonym to be identified by in the analysis of the study’s findings, and if they did not choose one, I assigned one to them. The majority of the participants were White and majored in psychology and/or criminology or sociology. To obtain this sample, recruitment of participants occurred through snowball sampling, which is a type of convenience sample. This sampling procedure allowed me to recruit participants within my own social network, participants whom I did not know, and those who were readily available to me to expand my sample. Additionally, I used SONA, a participant-recruitment tool, which allowed students to sign up to participate in the study and potentially receive course extra credit. These methods allowed me to gather data from Gen Z individuals currently attending a medium-sized Midwestern university.

Findings

Although I anticipated that social media would play a large role in advocacy and awareness of restorative justice, the findings did not support this idea. Instead, social media is more of a secondary source of information for restorative justice, rather than a primary source. Typically, educational settings such as the classroom served as a primary source of exposure to restorative justice. Additionally, I found that many participants had incomplete understandings of restorative justice and were unable to define or comprehend it fully. The vague understanding that participants did have of restorative justice was typically positive, however. Participants recognized the flaws of the current criminal justice system and generally critiqued the punitive approach to
dealing with crime. Participants also recognized the different demographic factors that might influence the American public’s perception of restorative justice, and they acknowledged how these demographics could create differences in opinion regarding restorative justice and other social justice issues. Many participants were hopeful about the future of restorative justice but lacked the capability to present tangible or substantial solutions to the issues the country faces regarding crime. Moreover, some participants saw restorative justice as a possible solution to mass incarceration and as a beneficial approach to dealing with offenders.

Awareness of Restorative Justice

One main takeaway from my research is that students at this small Midwestern university were not overly aware of or familiar with restorative justice and how these practices currently work. Six of my original 26 participants were excluded from analysis because of their lack of knowledge of restorative justice, and several more either had very vague understandings of restorative justice or confused restorative justice for a separate, but similar, social justice issue. Several participants \((n = 5; 25\%)\) mentioned that they were familiar with the idea of restorative justice but were unable to define it. Olive said, “I honestly probably couldn’t give you a direct definition of restorative justice, but I understand the concept of it.” Another participant, identified as Dixie, stated, “I wouldn’t be able to give you, like, an exact definition, but when I hear those on social media, like, or even just in general, like, I know what it’s about to refer to.” Similarly, Grey said, “I couldn’t like… I don’t… I couldn’t define them if you said, like, give me a definition.” All three of these responses were in response to the interview question “Are you familiar with either of the terms, ‘restorative justice’ or ‘community-based justice’?”

Participants believed that restorative justice is not a well-known practice and that without awareness, it might be hard to implement these practices successfully and with the support of the public. Blue said:

In order for more people to be in support, more people need to understand it and [the] depictions of it needs [sic] to change because… when I very first heard about [restorative justice], I didn’t know what the heck this was going to do. Obviously, the more you look into it, the more it makes sense, but I think most Americans will probably not be in favor right now.

Roxy said, “I think people just need to even know it exists, like first and foremost” regarding the potential widespread use of restorative justice programs. Billie suggested that “if they’re better at promoting it and [showing us] what they actually need, I’m pretty sure if they were to teach things that actually benefit a person, and we would see how an individual can change through this program, so just like promote it
more.” Furthermore, Dixie suggested that additional educational awareness is necessary for restorative justice programs to be implemented into society:

I hope that it will be drawn into more conversations, especially when it comes to courts of law, just kind of like working to educate people…. There’s like an opportunity for education, I hope, mostly like courts of law, shift their ideals and look at it from a more ethical standpoint, and I’m hoping ethics plays a larger role than vengeance will in the future.

Social Media

While it is true that social media is influential in shaping Gen Z perspectives (Duffett, 2017), more than half (n = 11; 55%) of participants reported that they had first learned of restorative justice in the classroom. Additionally, some participants (n = 3; 15%) later went to social media to supplement their initial understanding and opinions, and other participants reported that social media was a superficial source of information on restorative justice and other societal issue topics (n = 5; 25%). Violet said, “On the surface level, yes, they [social media platforms] have allowed me to gain a deeper understanding…. I don’t think they are as helpful as a class.” Similarly, in response to the same question, Ray said that media forms and platforms have “enhanced my understanding or supplemented the stuff that I’m learning about in class.” Both responses suggest that social media might play less of a direct role in restorative justice awareness for the public and that only those who are already familiar with restorative justice may be exposed to such messages.

The main platforms mentioned by participants included Facebook, TikTok, Instagram, Snapchat, Twitter, TedTalks, and Tumblr. Instagram was the most common platform mentioned by participants as a source of information regarding restorative justice. Twitter was the second most commonly mentioned social media platform, followed by TikTok. With regard to social media, participants frequently discussed the ease and efficiency of social media and how easy it was to access information. Olive said, “I think it’s the fast part of it, like it’s quicker and I feel like it’s short and to the point and I feel like I can get to the information faster…. I feel like you can get more information straight out of it through social media.” Similarly, Ray said, “Yeah, I think its accessibility” in response to why TedTalks and Instagram are influential and significant for messages of restorative justice.

Although participants reported that social media is an important part of their lives and that the convenience of social media is alluring, 10 participants (50%) reported that social media content can be misleading and unreliable. For example, Sally described how social media content is potentially specious:

I know it’s biased and it’s anything you see on the Internet. You have to take into account that that’s just one person’s point of view…. I know I have a
[mis]construed vision of what it even is just because I’m… just seeing it off of Snapchat, which is not a reputable source.

Similarly, Olive said, “I mean obviously they would think like the news and your [own, independent] research should be the best.” Olive’s views reflected most participants’ ($n = 12; 60\%$) sentiment that social media is a poor source of information for restorative justice and other topics and issues.

Participants indicated that it was sometimes difficult to distinguish between factual and fictional content on social media. Poppy said, “It’s sometimes hard, because you know you’re always told from when you’re… little that not everything on social media is real, which is true, but it kind of allows you to have all the information and kind of evaluate it for yourself to the best of your ability.” Poppy’s response suggests that some participants may see social media as a supplement to other forms of education about restorative justice. In other words, the onus is on the individual consumer to assess content in order to determine its validity, which requires an initial familiarity with the topic in order to discern between fact and fiction. This potential perception of social media may be contrasted with traditional forms of media, like newspapers, which have historically acted as initial contacts for consumers to learn more about current events and issues (Jones, 2006). Similarly, Blue said, “Given the nature of especially social media, with all of the opinions and all of the stuff that’s out there, and you know just because it’s on the Internet doesn’t mean it’s true or accurate.”

Similar to considering social media to be misleading, participants frequently viewed it as opinionated and biased as well. Interestingly, I did not directly ask participants if they viewed social media as unbiased or biased. Instead, these themes emerged without prompting as the interviews progressed. Most commonly, participants discussed social media as being biased and/or opinionated in a general sense, or politically polarized. Angel described their interactions and perceptions of using social media as a source of information regarding current events and other issues:

Social media will twist it to only give you like I said, usually, not a lot of middle ground. . . . I’ll see like on social media sometimes based on things that you see in the news they’ll be like stuff made based on either severely politically left or severe politically right based on their opinions.

These participants saw media portrayal of topics, including but not limited to restorative justice, as biased and therefore inaccurate. These participants displayed a lack of trust in different media sources, including news stations. Dixie said, “Personally, like if Fox News is posting on social media about something, I automatically know it’s biased.” Participants viewed content containing a strong
opinion as “biased” and often did not choose to further investigate the validity of the message.

Restorative Justice Use: Nonviolent versus Violent Crimes

Participants’ opinions and beliefs typically differed when participants were asked “How would you feel if someone was participating in a restorative justice program for a nonviolent offense?” and “How would you feel if someone was participating in a restorative justice program for a violent offense?” Most commonly, participants were more accepting of offenders who had committed nonviolent crimes, such as drug offenses, rather than of offenders who had committed violent offenses, such as armed robbery. For example, 15 (75%) participants conveyed viewing restorative justice as a potentially beneficial and positive method of dealing with nonviolent offenders, while 10 (50%) participants said the same for violent offenders. Ray said, “I think more serious or violent crimes or repeat offenders might be placed in more punishment-based facilities and then people who are maybe nonviolent offenders might fall to that restorative justice sort of spectrum.” Skylar said, “I would support it,” in reference to using restorative justice methods for nonviolent offenders. Skylar added, “Whenever it comes to drug crimes, I feel like direct punishment and incarceration isn’t really going to do anything… I feel like rehabilitation is a more promising solution to kind of handling [drug crimes].” The participants’ differing reactions to nonviolent and violent offenders suggest that initial uses of restorative justice could be aimed at nonviolent offenders before moving to violent offenders.

Some participants also believed restorative justice could help humanize offenders for the general public, and identify the root causes of crime. Jane said, “I think the more that we can view them [nonviolent offenders] as people who were in some maybe difficult circumstances and treat them as such and also looking at punishment more like treatment, that we may improve and not need traditional punishments anymore anyway.” Shay said, “I definitely think that makes more sense than putting them in jail,” in reference to using restorative justice programs for nonviolent offenders. Shay also said that violent offenders could benefit from restorative justice practices:

I think that would, I still think it would be better than a traditional form of punishment because you’re working at, like, the actual root of the problem. And, rather than just like reacting to the behavior you’re reacting to, like, what actually is causing the behavior.

Similarly, Lane said, “I almost feel that it’s [restorative justice] a lot more helpful for nonviolent offenses because I feel like it’s a lot easier to get down to root causes of something… and, I don’t know, try to fix the factors that cause nonviolent offenses.” Moreover, seven (35%) participants expressed apprehension about using restorative
justice practices rather than traditional forms of punishment with violent offenders. For example, Roxy said, “I would definitely support it, I would just be more cautious, I guess . . . just like if it was me as a community member or whatever, I would want maybe like a closer eye on them, I guess I would just like to have a tighter rein on them.” Participants also suggested implementing restorative justice practices into the current criminal justice system and traditional forms of punishment, especially for violent offenders, as a way to use caution with violent offenders and restorative justice programing. Dixie said:

I also think that there should be like, it should be like a stronger version, and there should also be limitations placed. They should get the education and it should be like a large part of what they’re doing. However, I feel like not necessarily prison but definitely like house arrest and things like that, because that’s dangerous to, like, society and a community, and when it comes to armed robbery, a lot of people could have been hurt, so I think, like, a combination of the two. I wouldn’t say I’m necessarily all for only restorative justice. I’d say I, like, part, I . . . I think parts of restorative justice are important to implement into other systems.

Additionally, I found that participants’ personal experiences with crime and offenders might affect their perceptions of restorative justice. One participant reported a personal experience with drugs and drug users and therefore said they would not be comfortable with drug crimes being directed toward restorative justice programs. This opinion was in the minority, as the majority of the participants were accepting of nonviolent crimes, including drug crimes, being directed toward restorative justice rather than punitive punishment. In addition, some participants recognized that their privilege as White college students influenced their perception of restorative justice and the criminal justice system, more broadly. In other words, participants recognized that the punitiveness of that criminal justice system disproportionately affects people of color, particularly low-income people of color.

Restorative Justice: Soft on Crime

When asked what participants thought the American public would think about restorative justice programs, many believed that Americans would view restorative justice as harmful or negative. Many believed that Americans would see restorative justice as a “second chance” or as being “soft” on crime. Olive said,

I think the fact that they might not perceive it as, like, fair, but they think these people are, if they commit any sort of crime, like they’re outwardly dangerous so . . . a lot of people don’t believe that people deserve second chances, and I kind of could see this as almost sort of a second chance.
Similarly, Skylar said, “I feel like some people may view restorative justice as though restorative justice isn’t truly effective or harsh enough,” and continued to suggest that Americans might see restorative justice as “the easy way out” of punishment. Liz said that Americans might see restorative justice as “not having to really pay punishment for it, like they’re just going to a program, it’s more like a slap on the wrist.” Additionally, restorative justice might be too abstract or intimidating to Americans, according to some participants. Violet said, “I would say the majority of Americans probably don’t like the idea of restorative justice programs. . . . I think that, unless I had a background in criminology and, like, sociology, something like victim-offender mediation would be too scary for a lot of people.” Moreover, Blue said, “I would say most Americans probably would not be in support just because, without really looking into it, some people will be like, it’s, like, an easier way out or it’s not enough.” Participants also suggested that Americans prefer traditional punishment because it is what the public is comfortable and acclimated with when it comes to dealing with crime. Ray said, “There’s still very much, like, this sentiment of ‘People need to pay for what they did and be punished for what they did’. . . . so I think restorative justice isn’t seen in the most positive light right now.” Shay echoed this sentiment: “I think sometimes people desire to see others punished or, like, held accountable, and they wouldn’t even see restorative justice as holding people accountable.” Dixie added, “A lot of people like relying on that prison system to feel safe,” which shows how the American people might be more accepting of things that feel like a tradition, or because it is how things have always been done. Participants saw the American public as wanting to avoid dealing with the possible risks of implementing something new, like restorative justice, into the punitive carceral system that has been historically used to deal with crime.

Factors That Influence Perceptions of Restorative Justice in America

While being asked to discuss their perceptions of how Americans view restorative justice, participants were also asked to theorize about potential factors that may influence Americans’ opinions. I asked participants about what specific demographic factors they thought would be influential in Americans’ perceptions of restorative justice, and the most common responses were political affiliation and race. Participants also mentioned age, wealth, and firsthand experiences with addiction or drug crime as factors that might influence how an American may perceive restorative justice. For analysis, I focus on political affiliation and race because these were the two demographic factors discussed extensively by participants. Both the frequency and extent of discussion regarding race and political affiliation suggest that participants believed these factors to be most influential in determining support for or opposition to restorative justice among the American public.
Most commonly, participants ($n = 12; 60\%$) mentioned how race influenced individuals’ perceptions of restorative justice. Additionally, many participants ($n = 11; 55\%$) discussed political affiliation as a demographic factor that might influence perceptions of restorative justice in the United States. Regarding political affiliation, participants typically discussed how they thought conservative-leaning Americans would be less likely to accept or agree with the implementation of restorative justice practices, compared to their liberal-leaning counterparts. Lane said, “Conservative people might feel more negatively towards restorative justice.” Two other participants succinctly articulated this belief:

Conservatives and Republicans are usually more on that punishment and, like, pay for what you did, whereas Liberals and Democrats are usually more of that, like, everyone can contribute to society and we should rehabilitate them and work together within the community. (Ray)

I actually think that more-left-leaning individuals would probably be more open to something along these lines, alternative routes, and I think more-conservative right-leaning individuals would probably, in my experience, be more supportive of the current system that we have. (Blue)

Political affiliation was frequently discussed in relation to possible demographic factors that might influence Americans’ perceptions of restorative justice. Skylar discussed how politics and political affiliation infiltrate many aspects of society and individual daily lives and activities:

I feel like almost everything in our life just boils down to political affiliation and it seems that, overall, you know, those on the liberal side of the spectrum seem to be more understanding and willing to work towards change, whereas those on the conservative end of the spectrum are just kind of… they don’t really want to give people a chance to change.

Participant responses regarding the differences between conservative- and liberal-leaning individuals and their perceptions on restorative justice were echoed frequently. Interestingly, some participants ($n = 3; 15\%$) saw political affiliation as a dividing demographic factor and a cause of conflict. Poppy said, “I feel like I’ve seen specific sides of politics being claimed as a specific movement or program and that kind of deter s the other political party from supporting.” Furthermore, Belle also indulged in the idea that the political parties are divided, saying, “The two parties are probably as divided as they’ve ever been, so I think they try and find everything to argue about, so if one party supported, the opposite party is probably gonna bash them.” Participants viewing the political parties as polarized might suggest that this division may pose an obstacle to the implementation of restorative justice practices because the two parties may not agree on the proposal. Shay, while discussing influencing
demographic factors and how Americans might value traditional punishment over restorative justice, said,

    I mean, since things have become increasingly polarized like the last few years, especially, I think that people have become very resistant to, like, hearing the opposing side of things or, like, wanting to challenge their beliefs, and restorative justice is like a challenge to the traditional narrative of punishment, so therefore people are… people who have endorsed the previous system would be resistant to that because it’s different and some people don’t like to change.

Similarly, the idea that Americans are resistant to change, or that Americans value traditions such as the punitive methods of the current criminal justice system, was discussed by one participant (5%):

    America holds, like, the prison system in really high standing, and it doesn’t really focus on reform or restoring or, like, trying to make you a better person; it’s just purely about punishing what you have done. I think it would be a lot more helpful, and I think they call it a reform system, but it’s not a reform system, and so I think it would be a lot more helpful if it actually was a reform system and, like, other countries have much higher rates of reform that have also, like, nicer jails and prisons and, like, it’s not like what we [America] do as a jail, it’s more of like—honestly, there’s some that look like hotels—it’s about, you know, learning from your mistakes, instead of just suffering from them. (Grey)

Participants like Grey saw the benefits of taking a rehabilitative approach to dealing with crime and offenders in America; however, many also acknowledged the deep-rooted value and tradition linked with the current punitive justice system and saw the connection between political affiliation as well.

    Race was another common demographic factor that participants in the study believed to be significant in an individual’s perception of restorative justice in America. Belle considered the stereotypes that exist within American society regarding crime and criminals and said, “Some people have stereotypes about what race is going to be committing a crime, so I think that kind of thing will play a role.” Participants also considered the difference in experiences with the criminal justice system between racial minorities and White people. Ray said, “I think people who aren’t White might view it a little bit more favorably just because they’ve probably had more experience with the system, whereas people who are White and don’t have that experience might not really take it all that seriously and might be more on that punishment side of things.” Grey also discussed the significance of police discrimination and sentencing disparities in American society and the impacts they might have on an individual’s perception of restorative justice: “People of color might
feel like they’re targeted by the police, and so people of color have, like, charges for 
opposition of marijuana and are serving time for ten times longer than any White 
person would, or a White person wouldn’t be jailed.” Similarly, Lane said, “I feel like 
minorities might go more for restorative justice because they’re unfairly represented 
and a disproportionate amount of people of color are in prisons, so I feel like they 
would prefer that [restorative justice] more than White people.” Similarly, Blue 
discussed how the current criminal justice system affects people of color differently 
than it does White people, often harmfully and destructively:

I also think race might definitely have a lot to do with it. Obviously, different 
races have different perceptions of the justice system, and, you know, for White 
people, the justice system has worked for a very long time, and it works great 
for them. But for people of color, the justice system doesn’t really work great for 
them a lot of times, so I think, you know, if it works for you, don’t change it, but 
if it doesn’t work for you, of course you’re gonna want to change it.

Interviews revealed that these participants recognized the existing racial stereotypes 
and discrimination within the criminal justice system in America, but their responses 
showed that they do not know how to fully solve these issues in a comprehensive 
approach.

Future of Restorative Justice

The final question of the interview dealt with what each participant saw as the 
future of restorative justice, and a potential balance between restorative justice and 
traditional punitive forms of punishment. Through interviews, the notion emerged 
that although the current justice system will never be completely abolished or 
reformed, restorative justice offers an additional way to treat and rehabilitate criminal 
offenders. Rainbow said, “Everyone’s still going to be in favor of traditional forms of 
punishment, just because that’s what we’ve known, but I feel like restorative justice 
could maybe start to balance it out.” Similarly, Jane said,

I don’t think that we will really ever get away completely from traditional forms 
of punishment from prison and probation, but I think that, again, moving 
forward, we’re just going to start to see more restorative justice-type programs 
to really start balancing out, you know, who really deserves to be locked up and 
who deserves a chance to show that they’re not a criminal.

However, the interviews also revealed that some participants \( n = 3; 15\% \) had doubts 
about restorative justice as a whole and saw it more as supplemental to traditional 
punitive forms of punishment for criminal offenders. Dixie acknowledged the 
practicality and value of restorative justice but also expressed reluctance:

I think it’s really important for society, but it also has its downsides because we 
don’t really like … personally, I don’t know if it’s, like, more effective than, like,
punishment in the way you know. But, I think it’s a really good idea. I think it’s, just like in a way, there needs to be some compromises between the two [restorative justice and the traditional justice system]. … I think parts of restorative justice are important to implement into other systems. … People aren’t ready to drift towards something new; they’d rather implement new things into it. However, the way I see it, it’s that the prison system needs to be completely reformed in order to even implement new things, because it’s so corrupt. So, people are like … they’re okay with adding change to the system that already exists, but they’re not ready to reroute.

In a similar sense, Grey mentioned that restorative justice needs to be properly funded in order to be successful, and that money may come from the current prison system at some point in the future:

I hope that the criminal justice system will defund the prison system a little bit more and put that money into programs. I don’t know if that will happen anytime soon. But I do hope that it does, and I think that’ll happen eventually. I just don’t know how soon and especially like private prisons, I think they might start closing more and more instead of having private prisons, they might do private reform programs.

Additionally, discussed how there should be specific and detailed descriptions of restorative justice, and guidelines and policies that depict when restorative justice methods are appropriate and when traditional forms of punishment should be utilized. Blue said,

If restorative justice doesn’t take a better grip, then I do think that the country will not phase out traditional justice anytime soon, if ever. I definitely think it would have to be more of a combined effort, you know, like I said earlier, more case by case, like meet these qualifiers, or, I don’t know, something like that. I think it would definitely even work in partnership, but I don’t see it ever overtaking traditional measures, though.

Overall, participants seemed hesitant to agree to abolishing the traditional justice system completely. Many suggested integrating the two and reforming the current system to be more rehabilitative in nature, rather than punitive. Belle said,

I think there’s gonna have to establish a really fine line about where the differences are and really identify so there’s, like, no room for debate when someone’s going through, like, trial or whatever. Like, they need to establish certain crimes, definitely get prison time, or what went into the crime, like if weapons were involved, so I think it just, like, would depend on what the person does.
Discussion

The American public has grown increasingly distrustful of the criminal justice system in recent history, especially of the punitiveness of the prison system in the United States. Recent events such as the murder of George Floyd and Ahmaud Arbery have spiked discussions related to social justice issues and brought the issue of policing and discrimination in the United States, especially on social media platforms, to the forefront of national discourse. Restorative justice offers a promising alternative to the current issues with the criminal justice system, such as mass incarceration and high recidivism rates (Morris & Maxwell, 2001; New York Times, 2021; Rossner, 2008; Umbreit, 2000; Van Ness & Strong, 2015). Through the completion of 20 interviews and the subsequent qualitative and emergent coding process, the research revealed several trends and patterns regarding the relationship between social media and perceptions of restorative justice. Although my findings suggest that social media does not play a significant role in informing this sample of Gen Z college students about what restorative justice is, it is possible that social media could be harnessed for restorative justice advocacy. Although research exists regarding restorative justice and social media (see Islam [2008] for exceptions), there is a gap in research at its intersection. Social media serves as a method of communication and connection; however, an unanticipated finding suggests that social media might not serve as a significant form of advocacy for restorative justice among students at a medium-sized Midwestern university.

Prior research with social media has revealed the power that media holds in society, and how it is used for marketing, advocacy, and personal agendas (Cheng, 2021; Funk, 2013; Islam, 2008; Wetstein, 1996). Restorative justice advocacy does not appear to be influential on social media platforms, however, given that the majority of college students in this study use social media regularly without knowingly observing or interacting with such content. Additionally, restorative justice content on social media platforms currently appears to serve as a supplemental tool, rather than a dominant or primary source of information or awareness. Instead, educational methods, such as the classroom, continued to be a recurring source of initial awareness of restorative justice for participants. Many participants mentioned that they were introduced to the idea of restorative justice through classwork and educational formats. The lack of exposure to restorative justice messages on social media suggests two possibilities for future research and direction: (1) social media should be harnessed more effectively to spread awareness of restorative justice to a wider audience and (2) educational settings should be utilized more often to teach about and expose people to restorative justice.

Some participants mentioned that without their background as a criminology/sociology/psychology majors or minors, they would not have been aware of restorative justice, let alone understand what it is or how it could be used in society.
These findings suggest that expanding the curriculum of other departments and fields could help increase awareness of restorative justice to people outside the typical fields of social justice.

Participants reported feeling that social media that contained strong opinions or emotions was “biased,” yet they did not verify if the information was indeed factual. The lack of fact-checking and the quick labeling of content as biased is dangerous because it perpetuates the notion that content that challenges an individual’s personal beliefs is biased or opinionated, regardless of its accuracy or validity.

This research revealed several possible patterns related to social media, crime, and forms of justice. For the most part, participants agreed on overall ideas and themes, such as the differences between nonviolent versus violent crimes. Most frequently, participants were more accepting of restorative justice for nonviolent offenders compared to for violent offenders. Many suggested that if restorative justice programs are used with violent offenders, extra caution and regulations should be applied. The participants’ hesitation to implement restorative justice programs with violent offenders might suggest that society start implementing restorative justice programs with nonviolent offenders before trying to use them with violent offenders. This might help ease society into the idea of a restorative approach to crime. Unsurprisingly, many participants believed that the public views restorative justice as “soft on crime.” The belief that restorative justice is an “easy way out” may stem from a lack of understanding of what restorative justice really is and what restorative processes truly entail. This research supports the idea that many people are not familiar with restorative justice, and this lack of familiarity likely fuels myths regarding restorative justice.

Importantly, many participants also recognized the flaws with the current criminal justice system, including recidivism rates and mass incarceration. Participants recognized that restorative justice could be an effective way to help reduce recidivism rates, mass incarceration, and bias within the criminal justice system, and this research suggests that overall public awareness of restorative justice should be increased. The findings also align with prior research regarding general public support for restorative justice practices for nonviolent offenders. Future research should focus on where restorative justice fits into violent offenses.

The findings also revealed trends related to demographic factors that influence the public’s perception of restorative justice, and how restorative justice could be used in society. Race and political affiliation were the two most frequently discussed demographic factors that participants believed to be influential in America’s perception of restorative justice. The interviews revealed that participants thought that conservative versus liberal political beliefs would be influential in how Americans perceive restorative justice. Specifically, the participants believed that conservative-leaning individuals would be less likely to accept restorative justice or to
view it in a positive light compared to more liberal-leaning individuals. Additionally, the participants believed that people of color would be more accepting of restorative justice, compared to White people, because of the history of bias, discrimination, and racism that exists within the current criminal justice system.

It is crucial that society address the root causes of crime, including poverty and institutional discrimination and racism, in order to effectively implement restorative justice practices. Restorative justice approaches often rely on community support and involvement; if the community in which an offender resides or commits a crime is unable to support and become involved with restorative justice attempts because of external factors (e.g. poverty, gun violence, lack of education), the entire community suffers and the program could be ineffective. The future of restorative justice relies heavily on mending the broken and symptomatic communities plagued with crime because of environmental and societal factors beyond their control.

Limitations

The participants I accessed through SONA systems and the snowball sampling methods limited my participant pool to mostly students majoring and/or minoring in psychology, criminology, and sociology. These areas of study may be more likely to address the criminal justice system, which might have made my participants more likely to know what restorative justice is, compared to the general population. This sample may have skewed beliefs regarding crime and the criminal justice system, as their studies focus on examining these issues. If the sample had included a broader spectrum of participants, it is likely that the overall perceptions of restorative justice would have been less positive and accepting. The participants in this study were enrolled in classes that examine the issues and structural failures of the criminal justice system, topics that had likely primed them to be accepting of reform in the criminal justice system. It is also important to note that all the participants in this study went to the same mid-sized Midwestern university. The lack of demographic and locational variety within participants likely also influenced the findings. Future research should aim to pull participants from different areas in the United States and people with varying demographic characteristics.
References


http://dlxs.richmond.edu/d/ddr/docs/papers/usesofnewspapers.pdf


https://www.pewresearch.org/ft_18-04-27_incarcerationrate_map/


INFLTRATING THE BOYS’ CLUB: THE PORTRAYAL OF WOMEN IN POPULAR CRIME DRAMAS

EMMA D. BERGQUIST, BUTLER UNIVERSITY
MENTOR: STEPHEN BARNARD

Abstract

The present study examines whether there has been a change in the portrayal of female criminal investigators over the past two decades. Cultivation theory suggests that media profoundly influences the views of the larger society, thus indicating significant implications of this portrayal. Five popular shows (Chicago P.D., Criminal Minds, CSI: Crime Scene Investigation, Hawaii Five-0, and Law & Order: SVU) were explored, with 200 episodes examined between the series using a mixed-methods approach. The main characters in each show’s first season and most recent season were coded and analyzed. Characters were coded for depth, authority, helplessness, and sexualization, among other variables. The percentage of women as main characters across the shows increased from 27.59% to 37.84% between 1998 and 2022. More specifically, 0% of characters who led the fictional investigative team in the first seasons were women, but 40% in the later seasons were women. These findings suggest that the representation of women in crime dramas has improved as the years have progressed, though women are still seen as subordinate to their male counterparts. The improvement in the representation of female investigators demonstrates to young girls that they can join the law enforcement profession and fosters more respect for the work of female investigators currently in the investigative field.

Introduction

Crime dramas are an increasingly popular form of media consumed by Americans as the collective interest in crime mysteries continues to grow (CineVue, 2020). These procedural shows, in which a team of detectives typically works together to catch the “bad guy,” made up half of the top 20 most-watched shows of 2022 (Porter, 2023). The entertainment of these shows is evident; however, the difference in representation between male and female characters in these shows gains much less notoriety.
Historically, male actors have been cast to play detectives, federal agents, and criminal investigators on television (Gates, 2006; King, 2008). Over the past few decades, however, many in American society have made a push for a more gender-equitable society (Aune & Redfern, 2010; Cox Han & Heldman, 2023). What is not known is whether this sentiment has translated into the representation of women in television, specifically in crime dramas.

Interest in crime shows has also led to an increased interest in careers in the criminal field (Barthe et al., 2013); however, the fictional portrayal of crime investigation on television is typically the only exposure to the field that the average person has access to (Barthe et al., 2013; Robbers, 2008). This illustrates the bigger implication of media on everyday life. It is evident that modern society, constantly surrounded by media, is influenced by its content. This phenomenon is demonstrated through the lens of cultivation theory, which will be discussed in depth later in this paper. Female representation is therefore an important factor in crime television shows. Representations in the media have real-world consequences, and equitable abilities of men and women must therefore be portrayed. Treating men and women equally on screen is important in furthering broader gender equality in society.

Gender inequality portrayed in the media and other aspects of the public sector can have detrimental effects on the society that we live in today. Studying social inequalities is important, but problematizing the persisting inequality in modern society is more vital. Many previous studies have emphasized the problematic nature of the inequitable social dynamics in society (Beddoes, 2011; David, 2008; Harrington, 2022; Mohajan & Mohajan, 2022; Weder, 2021). Based on this research, the present study operates under the notion that social inequalities, especially in terms of gender, must be understood as problematic and must be studied through a feminist lens.

More than two hundred years on from the first documented feminist text (Wollstonecraft, 1792), gender inequality still exists in our modern society. The Pew Research Center (Aragão, 2023) reported that, in 2022, on average, women earned 82% of the pay that men did, up just 2% from twenty years ago. Although there is tangible evidence of progress toward gender equality, American society has not yet reached equilibrium. Importantly, as many Americans continue to watch television, it is important to portray gender equality in this content so it may have a chance to translate into reality. The feminist goal of advocating for equality is the framework used in the present study.

Theories of Representation and Media Effects

As of 2022, just 23.2% of protective service positions in the United States, including those of police officers and detectives, were held by women (U.S. Bureau of
Labor Statistics, 2022). Although this may not be a surprising statistic, it is inequitable nonetheless.

Similarly to this real-world statistic, the media consumed by the American public perpetuates the idea that men dominate the criminal investigative field. The patriarchal framework of crime dramas may play a role in allowing gender inequality to persist in society beyond the screen.

Cultivation theory is the idea that long-term exposure to media shapes how an individual watching perceives the world (Perera, 2021). Simply stated, this theory asserts that prolonged exposure to the ideals shown on television can shape the social discourse of the person watching. Essentially, compared to those who spend less time consuming television content, those who spend more time in front of a television screen are more likely to accept and mirror in their own views the ideas they are exposed to on television (Morgan, 2009). One survey study found that television watchers who regularly viewed police procedural programs more readily believed in the traditional view of masculinity, including toughness and aggression, than did those who did not regularly view crime shows (Scharrer & Blackburn, 2018). This finding is consistent with the notion that male characters in crime dramas often engage in violence (Finger et al., 2010; Parrott & Parrott, 2015; Scharrer, 2012). The demonstrated effect of long-term exposure to television ideals is especially important to note given that the average American consumes 2 hours and 33 minutes of television per day, not including any time spent watching content on phones or tablets (Stoll, 2023). Through the lens of cultivation theory, it is clear that this amount of television watching is vital in creating the social perceptions of the viewing audience.

The present study hinges on cultivation theory’s assertion that the way people view the criminal justice system depends on how the system is shown on their television screens. As many media consumers likely do not have personal experience in detective rooms or governmental agencies, their knowledge is based on the characters they see in popular crime dramas. Because of this, consumers likely have an inequitable view of the investigative field, as male characters historically outnumber female characters in these shows (King, 2008). Women have long struggled to be given respect in criminal justice professions, and although there is still a long way to go, it is important that female characters be portrayed with the same respect as their male counterparts.

Though inequality in the sector of crime dramas may not seem like a large issue for society as a whole, the impact it has on viewers makes this an important topic to research. A relevant example is a phenomenon known as the CSI effect. This is the theory that individuals who view crime dramas place more emphasis on forensic evidence than and hold different attitudes toward the criminal justice system from those who don’t. Multiple studies have validated this theory, thus showing that a direct correlation exists between the television shows that people watch and their view of the
world (Hayes & Levett, 2013; Mancini, 2011; Tapscott, 2012). Though what is shown on screens might not be an accurate representation of reality, consumers still internalize the values of fictional portrayals. For example, multiple studies have found that viewers of Disney princess movies, which are explicitly fictional, are more likely to believe that men are more dominant in romantic relationships than women are (Coyne et al., 2016; Hefner & Kretz, 2021; Tonn, 2008). Such beliefs further perpetuate harmful gender stereotypes, something not exclusive to children’s films.

This unconscious bias against women in the media has been shown through the increasingly popular use of the Bechdel test. The test is a series of rules made popular by Alison Bechdel in a comic strip from 1985 to evaluate if a work of fiction is inclusive and representative of female characters. To pass the test, a film or show must follow three rules: (1) at least two women are featured, (2) these women have a conversation with each other, and (3) the conversation revolves around something other than a man (Friedman et al., 2017). This test distinguishes between a piece of media that merely includes women and a piece of media in which women have substance to their character. This test was used in evaluating the individual episodes of crime dramas in this study. Though women might be cast as main characters, they may not have the depth that their male counterparts do, especially in procedural crime shows.

Geena Davis, an actress herself, studied a group of films in 2015, finding that male characters received two times the screen time and two times the speaking time of female characters (Davis, 2015). More surprising from her findings is that even though men dominate the movie industry, films with women leads in 2015 grossed, on average, 15.8% more than films male leads (Davis, 2015). In other words, women can be just as, or more, successful in acting roles than men yet are not given nearly as many opportunities to do so. Based specifically on Davis’s work, the present study takes this idea a step further and quantitatively analyzes the time in which each gender is on screen without the opposite gender.

Patterns of Representation

Women have long been portrayed as subordinate to their male counterparts (Davidson, 2015; Rauch, 2012). Traditionally, women are seen as the homemakers and caregivers to their families while men are seen as the providers who go to work each day (Cohen, 2006; Gates, 2011; Smith & Granados, 2009). In many popular crime shows, the women on the team are often forced to choose between raising families and keeping their jobs (Griffin, 2017). Men in these shows are not forced to make this same decision. Moreover, women in crime shows who do have jobs are more likely to wear revealing clothing and be viewed as sexual objects in these shows than men who are meant to be in the same position (Davidson, 2015; Manion, 2015; Rauch, 2012). Women are also more likely to appear as victims of violence in popular investigative shows (Garland et al., 2017; Manion, 2015; Parrott & Parrott, 2015). Though
advancements have been made in the way women are represented on screen, crime dramas are still viewed as predominantly masculine shows (Garland et al., 2017; Griffin, 2017; Manion, 2015).

The men in popular crime shows are often considered strong, commanding, and able to balance their work and home lives (Griffin, 2017; Parrott & Parrott, 2015; Rauch, 2012; Scharrer & Blackburn, 2018). They often sexualize the women on their teams and consider themselves more important in their positions (Davidson, 2015; Griffin, 2017; Manion, 2015). Whereas most women are pursued by at least one of their coworkers in crime dramas, the men are not subjected to the same romantic plotlines (Griffin, 2017; Manion, 2015). This diminishes the women’s roles to being objects that the men want, instead of focusing on the work the women do on the team.

Although the quantity of representation is inarguably important, the quality of this representation is equally, if not more, important. The present study draws largely from Lindner and Barnard’s (2020) book All Media Are Social and its ideas about women in the media and the aspects of representation important to evaluate women in the media. First, numerical representation indicates how often members of a particular social group are shown in the media, the most obvious classification of representation. Women are underrepresented on screen and thus, even when represented well, are not showcased as much as their male counterparts. Second is the quality of representation, which is how women are shown in the roles that they do receive. Specifically, this evaluates whether women are depicted stereotypically, given authoritative roles, or portrayed in a sexualized way. Finally, the extent of centrality of representation indicates which gender dominates a certain field. In the case of crime dramas, men have historically dominated character representation (Gates, 2011; King, 2008). These aspects of representation are important to note in tandem with each other. Although representation may be thriving numerically, it is not positively progressing if these representations lack quality and centrality.

Effects of Representation

This study is based on the notion that it is important to represent female narratives equally to male narratives on screen, as female narratives have been overlooked for years. In the early days of television, shows focused on families in which the mother stayed home and the father went to work (Gates, 2011). Now, women are being portrayed in the media as being in the workforce yet are not given the same respect or high positions as men (Davidson, 2015). As life often imitates art, this is problematic for the modern world.

The formal theory for this phenomenon is called agenda setting, which is the notion that the media consumed by society influences what society thinks about (Sparks, 2015). Going a step further, the theory of framing points to the fact that media
can tell viewers how to think as well (Sparks, 2015). Used in conjunction, both theories highlight the impact on reality of what society views on television, making what is shown more important than just entertainment. Both agenda setting and framing have been replicated and validated over time and as media has become more culturally prominent in the everyday lives of most Americans (McCombs & Shaw, 1972; Vreese & Lecheler, 2016; Weaver, 2015). Importantly, the real-life effects of media have a profound impact on young, impressionable viewers, who tend to be limited in their ability to think abstractly to distinguish between fiction and reality (Perse & Lambe, 2016). In the context of the present study, this can hinder the advancement of women in investigative professions because there is little female representation in this field in popular media.

It is important to note that equality of status and resources may not always result in an equal proportion of men and women in a certain field. For example, even if women are able to join the investigative field with the same ease as men, they may decide not to, based on personal preference. However, the present study rests on the notion that these preferences may be based on the representation of women in the investigative field in the media. Through the lens of agenda setting and framing, it can be argued that young girls and women create their preferences in part because of what they consume through the media and that the portrayal of women in crime-fighting television shows can therefore directly influence their own decisions, even subconsciously, to join the investigative field.

As the world is slowly becoming a more gender-equal place, there is still a gender bias in choices of profession (Garland et al., 2017; Johansson et al., 2008; Ruble et al., 1984). Representations in the media may influence the jobs that young people choose to strive for. If women in law enforcement are seen in the media as weak or inferior, it stands to reason that fewer young girls will want to enter that field of work. Researchers have found that media representation and lack of role models is the predominant reason why young girls do not apply to certain professions (Garland et al., 2017; Johansson et al., 2008; Manion, 2015). Though directors and actors may not be consciously aware of the stereotypes that they are perpetuating, those stereotypes exist nonetheless and can be harmful to impressionable young viewers.

Research has shown that although men and women have found similar success when it comes to careers, there is a difference in how they are assumed to have gotten to their positions (Manion, 2015). While it is often assumed that men have worked their way up in their careers and have natural talent, when women are given authority, they are assumed to have been lucky or favored (Manion, 2015). Women in crime dramas are often thought of in the same way. Though they may have authority, it is not assumed that they worked for or deserve their positions as much as men would (Manion, 2015). In the present study, this type of sexual objectification of characters
is measured because a correlation has been found between sexualization in media and the self-objectification of women (Karsay et al., 2018).

The Present Study

Considering the aforementioned concerns, the present study aims to understand how the portrayal of women in crime television shows has changed since these shows began. In recent years, the country has become more aware of feminist issues in the media similar to this one, such as the TimesUp and MeToo movements. As a result, many popular broadcast companies and media outlets have made a conscious effort to portray women as equal to their male coworkers. By examining both the first and most recent seasons of popular crime dramas, researchers will be able to discern if this promise was fulfilled.

In a broad sense, this study asks the question “Does the portrayal of female law enforcement officers in popular crime dramas differ significantly from that of male law enforcement officers? Additionally, has the portrayal of female law enforcement officers changed throughout these shows from the first seasons to the most recent?”

Most previous literature on this topic examines film and television shows from several years ago (Davidson, 2015; Garland et al., 2017; Griffin, 2017); however, equality in society has progressed significantly since these findings were published. How much that progression has affected the media, however, has not been studied.

Specifically, in crime dramas, a genre growing rapidly in popularity, this push toward equality is important to portray on television. Based on the previous research mentioned and the changing climate of equality in the United States, the present study was created to test a variety of specific hypotheses:

- **H0**: Women will continue to be generally seen as subordinate to men in crime dramas in most recent seasons, though the gap between them will close.

- **H1**: Women will be less overtly sexualized by the men in the later seasons of these shows.

- **H2**: There will be a larger percentage of women playing regular roles on crime drama shows in their most recent seasons rather than their first seasons.

- **H3**: There will be a smaller disparity between men alone in a scene and women alone in a scene on the most recent seasons of crime drama shows.
Research Method

Research Design

The present study took methodological inspiration from Garland et al. (2017) and their analysis of federal agents in popular television crime dramas. As such, content analysis was used on 200 episodes of popular crime dramas. This methodology has been proven to be helpful in sociological research through the use of description, systemic identification, and quantification of themes, images, language, and ideas (Garland et al., 2017). Typically, content analysis is used to study written text, but it has also been used to analyze images in popular culture and draw themes from them. The present study used content analysis to analyze both the manifest and latent functions of the language, images, and scenes portrayed in the crime dramas studied.

Analyzing manifest content involves examining physically present and countable elements in the media, whereas analyzing latent content involves examining the symbolism or otherwise hidden meanings behind the manifest content (Berg, 2004, as cited in Garland et al., 2017). The study of latent content has been thought to be a less reliable form of analysis (Hesse-Biber & Leavy, 2006, as cited in Garland et al., 2017), though “it allows for more depth and greater validity in assessing information and is particularly helpful in examining social issues such as gender representation” (Garland et al., 2016, p. 8, as cited in Garland et al., 2017).

This study was a descriptive design, coded to be quantitative. This allows the researcher to analyze the frequency of interactions between seasons and characters. This creates clear and interpretable data across interactions in the shows. Moreover, there were qualitative analyses of specific quotes that were interpreted throughout the study. Each main character in the selected shows was coded and compared to the rest of the characters being studied.

Sample

To test the research hypotheses, five crime-based dramas were examined. These shows were selected by using the websites of major television networks (ABC, CBS, FOX, NBC), as well as viewership rankings identifying popular television shows (IMDb, indiewire.com, Chicago Tribune, The Guardian). Additionally, each show selected had at least one woman as a main character, involved a team of detectives, was on a mainstream television channel, and took place in the United States. The shows selected were Criminal Minds (CBS), Law and Order: Special Victims Unit (NBC), Hawaii Five-0 (CBS), CSI: Crime Scene Investigation (CBS), and Chicago P.D. (NBC). The shows were selected based on popularity, diversity in the years they aired, and the presence of a large investigative team. Each main character in these shows was coded
based on the traits displayed in episodes of the first and final seasons as well as the character’s interactions with other characters. A total of 4587 minutes and 55 seconds across the first seasons and 3979 minutes and 5 seconds across the final seasons were analyzed, for a total of 8567 minutes across 200 episodes of all five series. (See Appendix A.) Once the analysis of the original 200 episodes was complete, 10 randomly selected episodes from the originally coded set were reanalyzed to ensure the reliability of the coding. (See Appendix D.)

Coding

Before analysis of the first and last seasons of the selected television series, a pilot study was conducted to solidify the coding scheme to be used in the analysis. Using a random number generator, two episodes were selected from the entirety of each series. The primary researcher was the sole coder in the study to elicit consistent results throughout. The codebook (Table C1) for this study was solidified after the conclusion of a pilot study.

Characters examined in this study were only those classified as main characters. Main characters were defined as those who received individual credit in the opening credit sequence for the show. Chicago P.D. does not include an opening credit scene; the main characters were therefore defined as those played by actors named first in the credits and not classified as guest stars. Neither recurring characters nor guest stars were examined in this study. A character must have been classified as a main character for more than half of the season to be included in the analysis.

Based on the work of Garland et al. (2017), two coding instruments were used to analyze the characters in this study. The first codesheet examined each character individually. For this codesheet, the following characteristics were noted: marital status, parental status, experience level, whether they were saved, flirtatiousness, objectification, whether they were comforted, and whether they were interrupted during the episodes analyzed. (See Appendix C.)

The second coding sheet consisted of frequency data and compared men’s data to women’s data throughout each season. Variables included the genders of the main characters, the boss, the character saying the first line, the arresting officer(s), those involved in interrogations, the characters who had “Aha” moments, time each gender spent on screen without the opposite gender, characters with emotional attachment to a case, characters with revealing clothing, personal side characters, and whether the Bechdel test was passed in each episode. (See Appendix C.)

As a result of the initial pilot study, the researcher qualitatively analyzed dialogue between characters for references to sexuality or gender, typically traditional responsibilities for each gender, and discriminative statements. All of these were
noted and used for analysis with the other characters. Moreover, some qualitative comments were converted to quantitative data. For example, all instances of interruption, flirtatious behavior, comfort, objectification, correction, and saving were noted with a time stamp and examined for dialogue. Additionally, the number of episodes in which a character wore revealing clothing was noted. For clarification on code definitions, see Appendix C.

Results

As the analysis of the results from this study relies heavily on the coded definition of gender, it is important to explicitly note the definition used. For the purposes of simplifying gender comparison in the results, a character was gendered by the pronouns used to refer to them. All characters that were coded were referred to by the binomial pronouns of she and he in their respective shows; therefore, only the words men and women will be used in the analysis and discussion.

Twenty-nine main characters were analyzed across the first seasons of the examined shows; of those, eight were women (27.59%). More than half (57.14%) of characters were portrayed as married, divorced, widowed, or having a significant other, and fewer than half (27.59%) of the characters were portrayed as parents. There was one authoritative boss in each series (18.52%), and the majority of characters were portrayed as seasoned professionals (90.48%), while the minority were seen as new team members (18.52%). Of the 41 personal side characters introduced throughout the first season of the series, 80.49% were attributed to male characters while 19.51% were attributed to female characters.

In the final-season episodes analyzed, 37 main characters were coded. Of these, 37.84% were women, a slightly larger percentage of characters was portrayed as parents (29.73%), and 56.76% were portrayed as having, or having previously had, a significant other. As in the first seasons, there was one authoritative figure in each series (18.52%); however, a smaller number of characters was perceived to be new team members (8.11%). The 51 personal side characters introduced throughout the seasons were again attributed mostly to male characters (68.63%) and less to female characters (31.37%).

Character Depth and Private Life

As mentioned, 57.14% of characters were portrayed as currently or previously having a serious romantic partner during the first seasons of the shows analyzed. This includes only one female character (12.5%), however: Catherine Willows, who was in the process of getting divorced in the series CSI. In contrast, more than half of the male characters fit into this category (52.38%). The vast majority of female characters were portrayed as single (87.50%), while fewer than half of male characters were shown to
not have a current or previous romantic relationship (47.62%). Additionally, 12.5% of women had children, whereas 33.33% of men were fathers. The lone woman who did have a child during the first season of her show, Catherine Willows on CSI, was investigated by Child Protective Services for neglect (“Blood Drops”), based on her dedication to her job.

Although the fictional jobs in each show are demanding, seven male characters in the first seasons of the examined shows were able to successfully balance their home and work lives, while their female counterparts, such as Catherine, were portrayed as unable to do so. Male characters in the first season of each show were given more depth in their personal and professional lives, on average, than were female characters. This includes relationships, backstories, and families. Guest stars were more than four times more likely to be introduced as connected to a male character than to a female character (80.49%, compared to 19.51%).

There was also a slightly higher chance that a male character would have an emotional attachment to a case than would a female character (56.10% to 43.90%). This particular variable adds depth to the male characters and elicits an empathetic emotional attachment to the character from the audience. Interestingly, the disparity of this variable increased in the most recent seasons of these series. Of the 43 emotional attachments portrayed in the final seasons of the shows, the men accounted for 29 (67.44%), compared to the women’s 14 (32.56%).

In the most recent season of the studied shows, almost half of the women portrayed were shown to have a current or past serious relationship (42.86%), while a majority of men (65.22%) were portrayed as romantically involved presently or previously. Both men and women grew in this category. Furthermore, about a quarter of men and women were parents in the most recent seasons of their shows (30.43% and 28.57%, respectively). One parent, female officer Kim Burgess of Chicago P.D., did have to fight to keep custody of her daughter, reminiscent of Catherine Willows’s fight in the first season of CSI. Kim’s dedication to her job was also used against her in a child custody case, but she did end up remaining the legal guardian of her daughter (“Gone”).

Women in the most recent seasons of the five shows examined were portrayed as more competent than they previously were, yet they were still outnumbered and more strained than their male counterparts. In the most recent seasons examined, guest characters were still more than two times more likely (68.63% to 31.37%) to have personal connections to male characters than to female characters. Four women appeared in both the first and most recent seasons of their respective shows, and eight men did the same.
Authority and Autonomy

In each show, one character was portrayed as the boss or main authority figure on the team. In the first seasons, 100% of these authority figures were men. Additionally, 50.00% of women in the first seasons were portrayed as transfers or new members of the investigative team, while only 4.35% of men held the same ranking. Women were therefore more often portrayed as more naive and less knowledgeable than the men on the team. This is displayed in *Hawaii Five-0* (“Ho-apono”) when Kono, the lone female investigator, assumes she will be interviewing a young witness because she is a woman. Danny, her male colleague, responds to this by saying, “No, no, it’s not because you’re a woman. It’s because you’re a rookie, which is way worse.”

Female characters had been discounted as team members throughout the first seasons of the examined shows; sometimes even side characters did not realize the woman was a member of the team. This is especially apparent in *Criminal Minds* (“A Real Rain”). J.J., a member of the investigative team, travels to a journalist’s office to interview him for the case, and the man assumes she is not a federal agent. Small assumptions such as these undermine the authority and value of women on the team.

The male characters in crime dramas also feel as though they can make decisions for the women portrayed as working under them. Some of the dialogue used by male characters refers to women as objects rather than people who can make their own decisions. In *CSI* (“The Strip Strangler”), Gil Grissom was possessive of Sara Sidle, a first-year transfer investigator working under him, when she was asked to go undercover for the Federal Bureau of Investigation. Grissom said he “told Culpepper that he could not have Sara.” This quote clearly shows a possessive nature over a female member of the investigative team. Grissom is in charge of his unit and thus is given the ability to make decisions for women on his team. Similar comments are seldom used when referring to male characters. The autonomy of the women in the first seasons of these shows is greatly diminished by the men who limit them under the guise of protection.

Moreover, of the 106 episodes examined across the first seasons of the selected crime shows, the first line from a main character came from a woman only 24 times (22.64%). The officer arresting a perpetrator in the series was far less likely to be a woman, as well (15.87%). A woman was also unlikely to interrogate a suspect without a male colleague (3.39%), whereas a man often interrogated a suspect without a woman present (46.33%). This lack of independence for female investigators was most apparent in the time they spent on screen without male characters—only 154 minutes and 31 seconds (3.35%)—as compared to 1234 minutes and 28 seconds (26.73%) that male characters spent on screen without female characters. The sentiment is directly expressed in *Chicago P.D.* (“At Least It’s Justice”) when Erin
Lindsay is going to investigate with another female detective, Mia Sumners. Their boss, Hank Voight, stops them to say, “I think it’s better if partners are male-female.”

Female main characters were almost constantly supervised by their male colleagues and were far less likely to have the “Aha” moment in an episode when the case was broken wide open (32.54%). Depicting female characters as less competent than male characters was a pattern throughout the first seasons of the five shows analyzed, although there were improvements in the most recent seasons.

Two series, Criminal Minds and Law & Order: Special Victims Unit, promoted a woman to the authority figure in the department, moving the percentage of female authority figures to 40.00% and lowering the male percentage to 60.00%. New male team members in the final seasons remained similar, at 8.70%, while new female team members dropped to 20.43%. The final season of Criminal Minds takes this notion a step further, and Special Agent Emily Prentiss is offered a chance at becoming the FBI director (“Family Tree”), who then slates Special Agent Jennifer Jareau as her replacement for unit chief.

That is not to discount the sexism still apparent in these shows, such as when Sergeant Olivia Benson in Law & Order: Special Victims Unit is called a “bitter woman cop who’s spent her life getting back at men” (“And the Empire Strikes Back”). The prevalence and commonality of comments such as these across the shows studied can be viewed as a direct reflection of the public’s perception of female investigators.

Across the 94 episodes of the most recent seasons of the selected crime shows, the first line from a main character was still less likely to come from a woman (28.72%). Additionally, the “Aha” moments in episodes continued to be dominated by male characters (61.86%). The officers who made arrests in the series also continued to be far less likely to be women (23.58%). Women did have a higher chance of interrogating a suspect without male colleagues in the later seasons (19.49%); however, men did as well (51.30%). The gap between genders for the time they were portrayed alone in the episodes analyzed did shrink, with women alone 7.42% of the time and men alone 20.40% of the time, compared to 3.35% and 26.73%, respectively, yet a large disparity does still exist.

**Sexualization and Seduction**

The overt sexualization of female main characters was a large theme in the dialogue examined for this study. Revealing wardrobe was noted for both genders, with women, on average, wearing revealing clothing in 81.81% of episodes they appeared in across the first seasons. In contrast, men were shown in revealing clothing in just 17.50% of episodes they appeared in during the first seasons of the shows analyzed. Moreover, all but one of the female investigators were objectified during the first seasons of their shows (87.50%), and six were flirtatious throughout the first seasons.
(75.00%). This is a large percentage, especially compared to the 19.05% of male characters who were objectified and 33.33% who were portrayed as flirtatious.

Elle Greenaway in the show *Criminal Minds* was the last team member added to the existing squad of all men. She appeared in the first episode of the series as an undercover agent attempting to lure the perpetrator into a trap. An example of overt sexualization of female characters is shown when Elle’s male colleague Derek Morgan sexualizes her instead of congratulating her by saying, “Next time, show a little leg” (“Extreme Aggressor”).

Subtle comments such as these invalidate the female investigator’s work and attribute her success to her sexuality instead. Seven of the eight (87.50%) featured female characters of the first season were also romantically or sexually linked to a coworker. This manifested as a physical relationship or communicated crush from a man they worked with. Though these women were marketed as professionals, they still played the subordinate role of love interests to their male counterparts. Only 23.81% of men were in the same category. Kevin Atwater, an officer in *Chicago P.D.*, condemned his partner, Kim Burgess, for her romantic relationship with an officer in the intelligence unit, a higher division that they were both hoping to join one day, by saying, “If your big plan is to sleep your way into Intelligence, I’m not sure if that’s the right guy to make that happen” (“Turn the Light Off”).

The relationship between Kim Burgess and Adam Ruzek, the higher-ranking officer, ended up holding her back from a promotion. Kevin is promoted to the intelligence unit while Kim is not, and no punishment is ever handed to Adam for the same relationship (“A Beautiful Friendship”). This example perfectly demonstrates the double standard for male and female characters in popular crime shows. Although the female character is punished for engaging in a relationship with her coworker, the relationship is not directly addressed with the male character involved at all. The handling of this situation is a clear example of a woman being criticized for her sexuality while the man involved in the same transgression experiences no lasting impact.

Many women across the five analyzed shows faced sexualization and discrimination throughout their time in the first seasons of their respective shows through physical presentation as well. Two women were shown bent over, with camera angles coming from behind. Two more were shown in bikinis or underwear. Catherine Willows, of *CSI*, even wore a name tag saying “foxy” on her lab coat while in the morgue (“To Halve and to Hold”).

In the most recent season of the five series, the gap closed slightly between men and women in terms of revealing wardrobe (21.00% to 71.33%, respectively), though a large disparity still exists for this variable. There was one female investigator whose backside was filmed and commented on—a much lower percentage than in the first
seasons, yet still not zero. A smaller percentage of female detectives was objectified, at 21.43%, in the most recent seasons, compared to just 4.35% of men.

Female characters decreased significantly in flirtatiousness, with just 5 women (35.71%) categorized as doing so, comparable to 7 men (30.43%) in the same category; however, female characters continued to serve as romantic interests to their coworkers at a higher frequency than males. Nine of fourteen female main characters (64.29%) were involved with someone they worked with, while just six of twenty-three (26.09%) men were romantically involved with coworkers. This statistic insinuates that a woman must serve a purpose as a sexual object rather than solely being accepted as a professional.

Helplessness and Fragility

While the naive portrayal of many female characters is prevalent in many of the areas analyzed, there were three variables coded specifically to measure this characterization: instances of interruption, comfort, and saving. In the first season, 37.50% of the main female characters were interrupted, as opposed to 9.52% of male characters. Of the instances of interruption of women, 100.00% of the 8 instances were by men. Similarly, men interrupted men in four of the four (100.00%) instances that it occurred. Men thus were portrayed as the more dominant gender in these shows, as having the right for their voices to be heard over others. In the most recent season, 21.34% of women were interrupted, comparable to 21.74% of men. Of the three instances of women interrupted, two were by men (66.67%). Of the six instances that men were interrupted, five were by women (83.33%), a large difference from the first seasons.

Across the first seasons, 87.50% of women were comforted in some capacity. This excludes only one female character. Ten (90.91%) of these instances involved a male character being the comforter. In the first seasons, 9.52% of men were also comforted, just two characters, for a total of 3 instances, all of which involved a male character being the comforter. This insinuates that women need reassurance and comfort in a high-stress job whereas men can handle the job more easily. There was a large increase in the number of times that main characters were comforted from the first to the most recent seasons of the five shows. In the most recent seasons, male characters were comforted on 10 occasions and female characters were comforted on 13 occasions, with women comforting men 60% of the time and men comforting women 100% of the time. Proportionally, women still outranked men in being comforted, by 57.84% to 34.78%.

In the first seasons of the five examined shows, 62.50% of women were saved at some point in the season by a team member, as opposed to 42.86% of men being saved. Across these 18 instances, men were the saviors 16 times (88.89%). In the final
season, the gap shrank to 42.86% of female characters and 34.78% of male characters being saved. Impressively, of the 10 instances when a team member was the savior, women and men were equal, with 5 each (50.00%). Unexpectedly, men were more likely to need saving in both seasons, being the ones saved in 10 (50%) of the first-season situations and 8 (80%) of the most-recent-season situations.

Discussion and Conclusion

The present study examined the difference between male and female investigators in popular crime dramas. Specifically, it focused on the change in the treatment of the main characters in these shows over time. Four main hypotheses were tested. H0 predicted that women would continue to be seen as subordinate in the most recent seasons as well as the first seasons. This was generally supported, though the initial gap in the equality of men and women in these shows did shrink from the first to the last seasons. H1 predicted that women would be less overtly sexualized in the most recent seasons, including both in dress and discriminatory comments. Although this was partially supported, women continued to dominate the group of characters with revealing clothing. Interestingly, instances of men being sexualized increased although the proportion of women stayed relatively the same. This does add to equality in crime dramas and likewise points to a more sexualized society in the recent decade. H2 correctly predicted that more women would be in regular roles on the fictional investigative teams. As noted, findings indicated that there was a roughly 10% increase in female main characters from the first to most recent seasons of the shows examined: 27.59% to 37.84%, respectively. Finally, H3 predicted that women would be on screen alone more in later seasons than in the first seasons. This was proven true, though a large disparity continued to exist between the time that men and women spent alone on screen.

Similar to findings in Garland et al. (2017), female investigators were found to be overrepresented on the screen as opposed to in the real workforce. This is a relatively clear indication that shows are attempting to create a more equitable image of the profession on television. The present study also found, however, that stereotyping and misogyny do exist on the screen in popular crime dramas, albeit at a lower rate presently than in earlier seasons of the shows.

The culmination of the results of this study indicates that women continue to be sexualized and stereotyped in popular crime dramas, though at a decreased rate from before. This is most clearly shown in the finding that 87.50% of women in the first seasons were objectified, compared to 21.43% of women in the final seasons. This representation may be similar to the treatment of female investigators in real life, though it insinuates that this behavior is acceptable to anyone entering the investigative field in the real world.
In contrast, it is important to note the rise of women in positions of power. In the early seasons, men dominated authority across all the shows examined, suggesting to the public that men were superior to their female colleagues. In recent seasons, however, the number of men in the highest position on investigative teams decreased, with more women in charge, showing that both genders can lead effectively. Women were thus given more responsibility and were trusted more openly than they had been previously.

Moreover, the first season of almost every series examined included a woman as the newest addition to the team. This portrayed the women as naive, unintelligent, and less capable than the men on the team. Again, in more-recent seasons, both men and women are portrayed as experienced professionals as well as new hires, making another important move toward a more equitable distribution in the media’s portrayal of the criminal justice field.

Female characters were given more depth as a whole as the seasons progressed as was initially hypothesized; however, women remain the minority of the cast in crime dramas. More personal guest characters were introduced who were directly related to female characters in later seasons, yet this number was still fewer than those attributed to male characters. It is important to note, however, that more women had families and children in the later seasons of the shows, lessening the gap between men and women with responsibilities outside of work. Allowing this balance demonstrates that both men and women can handle both their work responsibilities and their personal responsibilities, something that was not shown before. As many main female characters in the first seasons were seen as support for their male counterparts, further seasons broadened their importance and independent relevance to the show. Women still outnumber men in being romantically involved with or the objects of affection of their male colleagues, however, at 64.29% to 26.09%.

In agreement with previous findings (Garland et al., 2017), on average, women were less likely than men to be portrayed as victims. In this study, it was recorded each time a main character was saved from danger by another main character. Interestingly, men were much more likely both to be saved and to be the ones saving their teammates. This finding opposes the traditional “damsel in distress” idea that women need to be saved. Women did require more comfort in the shows, however, especially from men, than did their male counterparts. These themes continued throughout the first and final seasons with little change in frequency. Both needing saving and needing comfort portray vulnerability, one physical and one emotional. The findings of this study fit the traditional stereotype that men are braver and women are weaker. While men were seen as more physically vulnerable and putting themselves in harm’s way, women were seen as more emotionally vulnerable and unable to cope with the challenges of the job.
Essentially, gender equality is slowly becoming more prevalent in popular crime dramas today. While there is still room to grow in this regard, many of the variables noted in this study found a much more gender-inclusive environment in the later seasons than when the shows first began. Many of the optimistic hypotheses formed before the present study was run were confirmed to an extent. While some shows examined were more male-dominated, such as Hawaii Five-0, they were met with increasingly female-dominated shows, such as Law & Order: Special Victims Unit. On average, female characters have been given more depth than they were previously allotted. The findings of this study indicate that gender equality in crime dramas is moving in the right direction.

Although the portrayal of investigators in the media is meant to be fictional, the implications that accompany them are real. The male-to-female ratio shown in the media may not currently be accurate, but it may influence the future of the profession. Cultivation theory indicates that the way women in the law enforcement industry are portrayed on television can translate to the real-world treatment of women in this field. As society attempts to create more-equitable environments in various job fields, including criminal justice, television shows are vital to that success. Shows willing to sacrifice the depth of female characters for the further development of male characters convey the idea that this is an acceptable practice in everyday life. Slow strides are being made in the crime drama genre to show equality between male and female colleagues, yet more needs to be done before the crime drama genre can truly be called inclusive.

Limitations and Future Research

The present study analyzed only the first and last seasons of the included shows; the progression of female representation therefore could not be fully understood on a more longitudinal scale. The timeline of these shows spans from 1998 to 2022, though it is important to note that there are, at times, sizable gaps in this timeline. The feminist movement has changed significantly since 1998, including the MeToo movement, changes in reproductive restrictions, and increased female representation in federal positions. Law & Order: SUV has so far spanned 24 years, airing as each of these historic changes happened; however, Chicago P.D. has been airing for only 8 years, long after important milestones such as the MeToo movement. Contextualizing each show individually and examining the impact of its historical timing may be valuable in future research. Whereas this study examined the beginning and end of the shows, the seasons in between could be studied to pinpoint when changes in gender representation for women occurred, how rapidly they changed, and what outside context may have influenced that change.

As a partly qualitative content analysis was used in this research study, some findings were subject to interpretation by the researcher; however, previous studies
have found that content analysis is an effective tool for examining media content, despite this subjectivity (Garland et al., 2017). The most prevalent limitation of this study is human error, including potential omissions of key dialogue or symbolism and stopwatch inaccuracies. The reanalysis of a subset of 10 episodes from the total sample showed that the subjective coding for each variable was reliable, however, as a comparison of the independent analyses using Krippendorff’s alpha (KALPHA) proved to be in an acceptable range. (See Appendix D.) Additionally, the focus of this study was on the main characters, rather than all characters, portrayed in crime dramas. Future research may be beneficial in addressing the portrayal of all women in these shows. Importantly, the representation of gender-nonconforming characters, as well as others who are categorized in minority groups, such as race or class, would be a significant way to further the research in this area.

This study was limited also in its genre and generalizability. Results may not be generalizable to film, written media, or other forms of storytelling. Moreover, genres outside of crime dramas may produce different results than those in this study. Additionally, the explicit implications of these results on the general public were not tested. Interviews with viewers or those involved in the making of crime dramas can aid in fully understanding the impact of the representation in these shows. With ample time, an analysis of participants’ views of criminal investigators before and after exposure to popular crime dramas would be important in creating a deeper understanding of the effect that media has on the general public. Further investigation into the effects of gender representation on society, such as these, can inform future projects of major media companies.
References


https://books.google.com/books?hl=en&lr=&id=vRSdBQAAQBAJ&oi=fnd&pg=PP1&dq=agenda+setting+media&ots=d_JuzfA8pH&sig=VO1SmM4XsdElKIA0-knatPOyTEU#v=onepage&q&f=false


https://psycnet.apa.org/record/2012-99110-448

https://core.ac.uk/download/5067341.pdf

https://doi.org/10.1002/9781118541555.wbiepc121

https://doi.org/10.1002/9781405186407.wbieca036.pub2


Appendix A

Table A1. Seasons and Episodes Examined

<table>
<thead>
<tr>
<th>Television series</th>
<th>Season #</th>
<th>Episode #s</th>
<th>Year(s) aired</th>
</tr>
</thead>
<tbody>
<tr>
<td>Law &amp; Order: SVU</td>
<td>1</td>
<td>1–22</td>
<td>1998–1999</td>
</tr>
<tr>
<td>Criminal Minds</td>
<td>1</td>
<td>1–22</td>
<td>2005–2006</td>
</tr>
<tr>
<td>Hawaii Five-0</td>
<td>1</td>
<td>1–24</td>
<td>2010–2011</td>
</tr>
<tr>
<td>Chicago P.D.</td>
<td>1</td>
<td>1–15</td>
<td>2014</td>
</tr>
<tr>
<td>Hawaii Five-0</td>
<td>10</td>
<td>219–240</td>
<td>2019–2020</td>
</tr>
<tr>
<td>Criminal Minds</td>
<td>15</td>
<td>315–324</td>
<td>2020</td>
</tr>
<tr>
<td>Chicago P.D.</td>
<td>9</td>
<td>165–186</td>
<td>2021–2022</td>
</tr>
<tr>
<td>Law &amp; Order: SVU</td>
<td>23</td>
<td>495–516</td>
<td>2021–2022</td>
</tr>
</tbody>
</table>

Appendix B

Table B1. Description of Variables of First-Season Coded Female Characters

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Distribution and descriptive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>Single = 87.50%; married = 0.00%; divorced = 12.50%; widowed = 0.00%; significant other = 0.00%</td>
</tr>
<tr>
<td>Parental status</td>
<td>No = 87.50%; yes = 12.50%</td>
</tr>
<tr>
<td>Experience</td>
<td>New = 50.00%; seasoned = 50.00%</td>
</tr>
<tr>
<td>Saved</td>
<td>No = 37.50%; yes = 62.50%</td>
</tr>
<tr>
<td>Flirtatious</td>
<td>No = 25.00%; yes = 75.00%</td>
</tr>
<tr>
<td>Objectified</td>
<td>No = 12.50%; yes = 87.50%</td>
</tr>
</tbody>
</table>
Table B2. Description of Variables of First-Season Coded Male Characters

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Distribution and descriptive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>Single = 47.62%; married = 19.05%;</td>
</tr>
<tr>
<td></td>
<td>divorced = 14.29%; widowed = 9.52%;</td>
</tr>
<tr>
<td></td>
<td>significant other = 9.52%</td>
</tr>
<tr>
<td>Parental status</td>
<td>No = 66.67%; yes = 33.33%</td>
</tr>
<tr>
<td>Experience</td>
<td>New = 4.35%; seasoned = 95.65%</td>
</tr>
<tr>
<td>Saved</td>
<td>No = 57.14%; yes = 42.86%</td>
</tr>
<tr>
<td>Flirtatious</td>
<td>No = 66.67%; yes = 33.33%</td>
</tr>
<tr>
<td>Objectified</td>
<td>No = 80.95%; yes = 19.05%</td>
</tr>
<tr>
<td>Comforted</td>
<td>No = 90.48%; yes = 9.52%</td>
</tr>
<tr>
<td>Interrupted</td>
<td>No = 90.48%; yes = 9.52%</td>
</tr>
</tbody>
</table>
### Table B3. Description of Variables of Last-Season Coded Female Characters

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Distribution and descriptive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>Single = 57.43%; married = 14.29%;</td>
</tr>
<tr>
<td></td>
<td>divorced = 7.14%; widowed = 0.00%;</td>
</tr>
<tr>
<td></td>
<td>significant other = 21.43%</td>
</tr>
<tr>
<td>Parental Status</td>
<td>No = 71.43%; yes = 28.57%</td>
</tr>
<tr>
<td>Experience</td>
<td>New = 21.43%; seasoned = 80.57%</td>
</tr>
<tr>
<td>Saved</td>
<td>No = 57.14%; yes = 42.86%</td>
</tr>
<tr>
<td>Flirtatious</td>
<td>No = 65.29%; yes = 34.71%</td>
</tr>
<tr>
<td>Objectified</td>
<td>No = 78.57%; yes = 21.43%</td>
</tr>
<tr>
<td>Comforted</td>
<td>No = 42.86%; yes = 57.84%</td>
</tr>
<tr>
<td>Interrupted</td>
<td>No = 78.57%; yes = 21.43%</td>
</tr>
</tbody>
</table>

### Table B4. Description of Variables of Last-Season Coded Male Characters

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Distribution and descriptive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>Single = 34.78%; married = 26.09%;</td>
</tr>
<tr>
<td></td>
<td>divorced = 4.35%; widowed = 4.35%;</td>
</tr>
<tr>
<td></td>
<td>significant other = 30.43%</td>
</tr>
<tr>
<td>Parental status</td>
<td>No = 69.57%; yes = 30.43%</td>
</tr>
<tr>
<td>Experience</td>
<td>New = 8.70%; seasoned = 91.30%</td>
</tr>
<tr>
<td>Saved</td>
<td>No = 65.22%; yes = 34.78%</td>
</tr>
<tr>
<td>Variable and coding</td>
<td>Distribution and descriptive findings</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Gender (N = 29)</td>
<td>Man = 72.41%; woman = 27.59%</td>
</tr>
<tr>
<td>Boss (N = 5)</td>
<td>Man = 100.00%; woman = 0.00%</td>
</tr>
<tr>
<td>First line (N = 106)</td>
<td>Man = 77.36%; woman = 22.64%</td>
</tr>
<tr>
<td>“Aha” (N = 126)</td>
<td>Man = 67.46%; woman = 32.54%</td>
</tr>
<tr>
<td>Arresting officer (N = 126)</td>
<td>Man = 84.13%; woman = 15.87%</td>
</tr>
<tr>
<td>Interrogation (N = 177)</td>
<td>Men = 46.33%; women = 3.39%; both = 50.28%</td>
</tr>
<tr>
<td>Emotional (N = 41)</td>
<td>Men = 56.10%; woman = 43.90%</td>
</tr>
<tr>
<td>Wardrobe</td>
<td>Man = 17.50%; woman = 81.31%</td>
</tr>
<tr>
<td>Time alone (N = 4617:55)</td>
<td>Men = 26.73%; women = 3.35%</td>
</tr>
<tr>
<td>Personal side characters (N = 41)</td>
<td>Men = 80.49%; women = 19.51%</td>
</tr>
<tr>
<td>Bechdel test (N = 106)</td>
<td>Failed = 77.36%; passed = 22.64%</td>
</tr>
</tbody>
</table>
Table B6. Last-Season Description of Gendered Variables

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Distribution and descriptive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (N = 37)</td>
<td>Man = 62.16%; woman = 37.84%</td>
</tr>
<tr>
<td>Boss (N = 5)</td>
<td>Man = 60.00%; woman = 40.00%</td>
</tr>
<tr>
<td>First line (N = 94)</td>
<td>Man = 71.28%; woman = 28.72%</td>
</tr>
<tr>
<td>“Aha” (N = 97)</td>
<td>Man = 61.86%; woman = 38.14%</td>
</tr>
<tr>
<td>Arresting officer (N = 106)</td>
<td>Man = 76.42%; woman = 23.58%</td>
</tr>
<tr>
<td>Interrogation (N = 154)</td>
<td>Men = 51.30%; women = 19.49%; both = 29.22%</td>
</tr>
<tr>
<td>Emotional (N = 43)</td>
<td>Man = 67.44%; woman = 32.56%</td>
</tr>
<tr>
<td>Wardrobe</td>
<td>Man = 21.00%; woman = 71.33%</td>
</tr>
<tr>
<td>Time alone (N = 3679:05)</td>
<td>Men = 20.40%; women = 7.42%</td>
</tr>
<tr>
<td>Personal side characters (N = 51)</td>
<td>Men = 68.63%; women = 31.37%</td>
</tr>
<tr>
<td>Bechdel test (N = 94)</td>
<td>Failed = 68.09%; passed = 31.91%</td>
</tr>
</tbody>
</table>
Table B7. First- and Last-Season Coded Female Characters

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Distribution and descriptive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women on team</td>
<td>First = 27.59%; last = 37.84%</td>
</tr>
<tr>
<td>Female boss</td>
<td>First = 0.00%; last = 40.00%</td>
</tr>
<tr>
<td>Female first line</td>
<td>First = 22.64%; last = 28.72%</td>
</tr>
<tr>
<td>Female “Aha”</td>
<td>First = 32.54%; last = 38.14%</td>
</tr>
<tr>
<td>Female arresting officer</td>
<td>First = 15.87%; last = 23.58%</td>
</tr>
<tr>
<td>All female interrogation</td>
<td>First = 3.39%; last = 19.49%</td>
</tr>
<tr>
<td>Female emotional</td>
<td>First = 43.90%; last = 32.56%</td>
</tr>
<tr>
<td>Female wardrobe</td>
<td>First = 81.31%; last = 71.33%</td>
</tr>
<tr>
<td>Women time alone</td>
<td>First = 3.35%; last = 7.42%</td>
</tr>
<tr>
<td>Personal side characters for women</td>
<td>First = 19.51%; last = 31.37%</td>
</tr>
<tr>
<td>Passed Bechdel test</td>
<td>First = 22.64%; last = 31.91%</td>
</tr>
</tbody>
</table>
Table B8. First- and Last-Season Coded Male Characters

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Distribution and descriptive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men on team</td>
<td>First = 72.41%; last = 62.16%</td>
</tr>
<tr>
<td>Male boss</td>
<td>First = 100.00%; last = 60.00%</td>
</tr>
<tr>
<td>Male first line</td>
<td>First = 77.36%; last = 71.28%</td>
</tr>
<tr>
<td>Male “Aha”</td>
<td>First = 67.46%; last = 61.86%</td>
</tr>
<tr>
<td>Male arresting officer</td>
<td>First = 84.13%; last = 76.42%</td>
</tr>
<tr>
<td>All male interrogation</td>
<td>First = 46.33%; last = 51.30%</td>
</tr>
<tr>
<td>Male emotional</td>
<td>First = 56.10%; last = 67.44%</td>
</tr>
<tr>
<td>Male wardrobe</td>
<td>First = 17.50%; last = 21.00%</td>
</tr>
<tr>
<td>Men time alone</td>
<td>First = 26.73%; last = 20.40%</td>
</tr>
<tr>
<td>Personal side characters for men</td>
<td>First = 80.49%; last = 68.63%</td>
</tr>
</tbody>
</table>

Appendix C

Table C1. Master Codebook

<table>
<thead>
<tr>
<th>Code</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boss</td>
<td>The character who is making the decisions for the rest of the team or is referred to as a higher authority than other</td>
<td>Prentiss (to JJ): “If I were to move on from my position, I had you in mind to take over as unit chief” (CM, 15.8).</td>
</tr>
<tr>
<td></td>
<td>characters</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Parental status</td>
<td>A child of a main character mentioned or shown in the show</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Catherine: “My job does take me away from my daughter” (CSI, 1.10).</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td>A significant other or ex-significant other of a main character mentioned or shown in the show</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stabler goes home and talks to his wife about the case he is working on (SVU, 1.18).</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>The amount of time that a main character had spent on the fictional investigative team. A new member could be a new transfer or new hire, while an experienced member had an established role on the team.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Man: “How you been, Elliot?” Stabler: “Good, everything’s good. This is my new partner, Olivia Benson.” (SVU, 1.2)</td>
<td></td>
</tr>
<tr>
<td>Wardrobe</td>
<td>An outfit that shows the skin of a main character below the collarbone or above the knee</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Garcia wears a shirt that shows her cleavage (CM, 15.9).</td>
<td></td>
</tr>
<tr>
<td>Emotional</td>
<td>A main character is emotionally attached to the case and becomes overly involved with the victim or investigation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sara: “You want to sleep with me? … That way when I wake up in a cold sweat under the blanket hearing [the victim] Kaye’s screams, you can tell me it’s nothing. It’s just empathy.” (CSI 1.10)</td>
<td></td>
</tr>
</tbody>
</table>
| Corrected | A main character is proven wrong or corrected by another character | Nick: “She was stabbed, at least a dozen times—a screwdriver, maybe . . . like a spike.”
Grissom: “No. The gouges on her ribs are unusual. The instrument had to be slightly curved.” (CSI, 1.6) |
| Saved | A main character would have died or been injured if another character had not come to their rescue | A suspect has a gun pointed at Adam, and Quinn shoots the suspect before he can kill Adam (HF0, 10.21). |
| Flirtatious | Comments a main character makes to create a romantic relationship with another character, or comments made to sexualize another character | Garcia: “If I figure it out, does it earn me a night of passionate lovemaking?”
Morgan: “Most definitely sweetness.” (CM, 1.3) |
<p>| Romantic | A main character is portrayed as the object of affection for someone they work with | Reid asks J.J. to go on a date with him (CM, 1.4). |
| Objectified | A comment is made toward a main character about their body or sexuality | Man (to Lindsay): “Why don’t you bring that ass over to my place sometime, girl. I’ll let you in.” (CPD, 1.1) |
| <strong>Comforted</strong> | A main character going through an emotional episode is consoled by another main character | Quinn is crying, and Adam puts his arm around her to console her (HFO, 10.19). |
| <strong>Interrupted</strong> | One character is spoken over before they can finish their thought or sentence | Elle: “And a blow that hard wouldn’t scare her. It’d probably just—” Morgan: “Probably knock her unconscious.” (CM, 1.15) |
| <strong>Gender</strong> | The pronouns that a main character is referred to by | Warrick: “No, you’re just siding with Sara ’cause you got a crush on her” (CSI, 1.17). |
| <strong>Bechdel</strong> | Two or more named female characters have a conversation without a man about something other than a man | Benson talked to a female victim about her daughter (SVU, 23.13). |
| <strong>First line</strong> | The first main character who has a speaking line in the episode | After an opening scene involving guest stars, Benson speaks to her son: “I said no phones at the table” (SVU 23.5). |
| <strong>Arresting officer</strong> | The main character who arrested the offender | Danny puts handcuffs on the suspect after finding him (HFO, 1.18). |
| <strong>Male-female comments</strong> | Objectifying or flirtatious comments made by a male main character to a female character | Morgan (to Elle): “Next time, show a little leg.” (CM, 1.1) |
| <strong>Female-male comments</strong> | Objectifying or flirtatious comments made by a female main character to a male | Female lab tech (to Brass): “You want to be different? Pin me up against a wall; lay one on me like you mean it.” |</p>
<table>
<thead>
<tr>
<th>Character</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decoy/Lure</td>
<td>A main character is used as bait in an undercover operation</td>
<td>Burgess is used as bait for thieves in a park to steal from as all other members of the team watch to catch the offender (CPD, 1.9).</td>
</tr>
<tr>
<td>Interrogation</td>
<td>A main character is present in the room when a suspect is being interrogated</td>
<td>McGarrett and Danny are the only investigators in the interrogation room asking questions to a suspect (HFO, 1.5).</td>
</tr>
<tr>
<td>Aha</td>
<td>The moment in the show when a main character puts the pieces of the case together, which leads to the capture of the offender</td>
<td>Reid finds a CD hidden in a suspect’s room that reveals the password to discovering where the victim is hidden, leading to the victim’s rescue (CM, 1.1).</td>
</tr>
<tr>
<td>Women alone</td>
<td>One or more main female character(s) are on the screen without male characters with speaking roles (excludes talking on the phone to a male character)</td>
<td>Benson and the female district attorney have a private conversation alone in an elevator (SVU, 23.8).</td>
</tr>
<tr>
<td>Men alone</td>
<td>One or more main male character(s) are on the screen without female characters with speaking roles (excludes talking on the phone to a female character)</td>
<td>Halsted talks to a male veteran alone outside of a veterans’ meeting (CPD, 9.3).</td>
</tr>
</tbody>
</table>
### Personal Side Characters

A guest character is introduced as having a previously formed relationship with a main character

Danny: “Is that your sister in the car?”

McGarrett: “Yeah, I just came straight from the airport.” (HFO, 1.4)

---

### Appendix D

#### Table D1. Reliability Recoding

<table>
<thead>
<tr>
<th>Variable and coding</th>
<th>Finding used in study</th>
<th>Recoded findings</th>
<th>KALPHA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bechdel</td>
<td>Pass = 3</td>
<td>Pass = 4</td>
<td>0.791</td>
</tr>
<tr>
<td></td>
<td>Fail = 7</td>
<td>Fail = 6</td>
<td></td>
</tr>
<tr>
<td>First line</td>
<td>Man = 7</td>
<td>Man = 7</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Women = 3</td>
<td>Women = 3</td>
<td></td>
</tr>
<tr>
<td>Aha</td>
<td>Men = 9</td>
<td>Men = 8</td>
<td>0.919</td>
</tr>
<tr>
<td></td>
<td>Women = 1</td>
<td>Women = 1</td>
<td></td>
</tr>
<tr>
<td>Arrests</td>
<td>Men = 11</td>
<td>Men = 14</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>Women = 1</td>
<td>Women = 1</td>
<td>1.000</td>
</tr>
<tr>
<td>Interrupt/Ignore</td>
<td>Men = 0</td>
<td>Men = 0</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Women = 0</td>
<td>Women = 0</td>
<td></td>
</tr>
<tr>
<td>Decoy/Lure</td>
<td>Men = 1</td>
<td>Men = 1</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Women = 0</td>
<td>Women = 0</td>
<td>1.000</td>
</tr>
<tr>
<td>Interrogation</td>
<td>Men = 6</td>
<td>Men = 7</td>
<td>0.978</td>
</tr>
<tr>
<td></td>
<td>Women = 0</td>
<td>Women = 0</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Both = 3</td>
<td>Both = 3</td>
<td>1.000</td>
</tr>
<tr>
<td>Time alone</td>
<td>Men = 136:03</td>
<td>Men = 138:05</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>Women = 13:58</td>
<td>Women = 13:51</td>
<td>0.731</td>
</tr>
<tr>
<td>Wardrobe</td>
<td>Men = 7</td>
<td>Men = 7</td>
<td>0.791</td>
</tr>
<tr>
<td></td>
<td>Women = 13</td>
<td>Women = 13</td>
<td>0.797</td>
</tr>
<tr>
<td>Flirt</td>
<td>Men = 1</td>
<td>Men = 2</td>
<td>0.967</td>
</tr>
<tr>
<td></td>
<td>Women = 0</td>
<td>Women = 0</td>
<td>1.000</td>
</tr>
<tr>
<td>Saved</td>
<td>Men = 2</td>
<td>Men = 3</td>
<td>0.747</td>
</tr>
<tr>
<td></td>
<td>Women = 0</td>
<td>Women = 0</td>
<td>1.000</td>
</tr>
<tr>
<td>Comforted</td>
<td>Men = 1</td>
<td>Men = 2</td>
<td>0.967</td>
</tr>
<tr>
<td></td>
<td>Women = 1</td>
<td>Women = 1</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Men = 0</td>
<td>Women = 0</td>
<td>Men = 0</td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>Objectified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women = 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men = 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women = 0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
POPULISM: A COMPARATIVE STUDY OF THE RISE OF THE ALT-RIGHT IN GERMANY AND THE UNITED STATES

KATE O. TOBIN, BUTLER UNIVERSITY
MENTOR: MELISSA ETZLER

Abstract

The far-right extremist movements in Germany and the United States have gained attention and proved concerning, manifesting in the forms of terrorism, nationalism, and xenophobia. The radical right often utilizes populism to undermine liberal institutions, with a proneness to discrimination and violence. Right-wing populism in Germany has been a movement founded on fear, which has expanded to include governmental and political institutions. In the United States, populists often criticize the political elite for catering to the needs of minorities. The influence of the economic crisis on the lower-income, blue-collar areas of the country has been impactful enough to allow for populist rhetoric to gain a foothold in America. Populism is successful only if accompanied by effective propaganda and rhetoric tactics. In Germany, far-right populist parties focus on the political elites’ disrespect for German nationalism. In the United States, Donald Trump was able to push rhetoric through an unconventional campaign and build a following that will continue to grow quietly, despite the ending of his term of presidency, through hateful ideologies that portend a cataclysmic event. This event may be necessary for far-right wing extremism to be collectively rejected from American society with as much enthusiasm as seen with the rejection of Nazism following WWII and the Holocaust.

Leading up to and during Donald Trump’s time in office, and thereafter, comparisons between Trump and Adolf Hitler have become commonplace in American political discourse, used to characterize Trump as a leader. Various media platforms and individuals of influence have publicly commented on the likeness. *Saturday Night Live* has incorporated the comparison into its skits, as have other late-night comedians. Hosts of *The View* also noted concerning parallels leading up to the general election. Glenn Beck, a previous host of Fox News, made the observation while appearing on ABC’s *This Week* segment (Friedman, 2016). This statement has been used to criticize Trump’s presidency and claim that the ideologies he promoted during his tenure are fascist in nature. Although many may agree that Trump has become a right-wing extremist enabler, the notion that Trump is comparable to Hitler runs the risk of being based solely on the assumption that these two leaders align because they
both belong to far-right groups. This general claim frequently does not derive from an actual comparative analysis of what was portrayed by their campaigns, leadership, and rhetoric. This is not to say that there are not similarities. There are numerous disturbing ways in which Trump has emulated elements of Hitler’s campaigns and propaganda. Additionally, there are similarities in how the ideologies of their parties have manifested, as well as a consistent issue of violence and hatred present within their agendas and their supporters. This paper aims to perform a more detailed analysis of the two leaders, taking into consideration the political and economic climates during their candidacies, their end goals, the foundations of their campaigns, and, most prominently, how populism and its relation to nationalism played a role in their propaganda and rhetorical tactics. There are contrasts in their leadership styles that are difficult to acknowledge without a clear understanding of populism. The propaganda and rhetoric that have been employed require a deeper look for accurate analysis, comparison, and contrast of the two leaders. These tools also help us compare the modern far-right wing movements seen in both Germany and the United States and speculate what this may mean going forward.

This paper has its genesis in a project that I began in 2019 on Nazi propaganda and the rhetoric that was able to convince an entire population to turn against its European neighbors. I could not have imagined that, just a few years later, the political tide would turn in such a way that it would bring these ideas back to the forefront of my mind, as well as society’s. Scapegoating, racism, “othering”—these are terms with which the public is confronted almost daily. This paper seeks to further explore the ideology of populism and its overlap with nationalism, using this analysis to compare the leadership styles of Donald Trump and Adolf Hitler, as well as populism’s influence in American and German politics today. This will lead to the conclusion involving a more detailed analysis of the existence of the alt-right in the United States today, and how we ultimately may witness a violent, cataclysmic ending to this extremist movement.

The first section of this paper will include a literature review on the ideology of populism, including descriptive characteristics that can aid in its recognition. These characteristics include the populist figure as the “voice of the people,” anti-elitist beliefs, scientific skepticism, friend-enemy distinctions, protectionism, and scapegoating. This portion of my work will also seek to compare populism and nationalism, including where they overlap and where they differ. From there, I look at Hitler as a nationalist leader who at times explored populist propaganda tactics. This section will also include political developments since WWII that have contributed to populism’s rise in Germany in more recent decades. This portion of my work will end with an analysis of today’s right-wing Alternative für Deutschland (AfD) party, including the group’s overlap with populism and the group’s effective use of rhetoric. I then highlight some of the instances when populism has been present in American politics throughout history and leading up to the 2016 presidential election. An
examination of Donald Trump as a populist follows, including a more in-depth exploration into his scapegoating and ostracizing of certain identifiable groups. Finally, I argue that Hitler, despite at times implementing populist rhetorical devices, focused on a more nationalist appeal. Trump, however, can be categorized as a populist who harbors an appeal to nationalist sympathizers. It is the dissonance between Trump’s intentions and his base’s underlying anger that will provide the foundation for an explosion of extremism.

My hope for this paper is that its readers will be able to compare Hitler and Trump more accurately, acknowledging similarities found in their propaganda and rhetoric while also being aware of the men’s ideological differences. I aim for this work to put this common comparison in perspective and to leave readers more educated on the similarities. In addition, this work should enlighten individuals on the power of populism in the political landscape today, and my breakdown of the ideology will, I hope, allow for readers to recognize when populism is present in a campaign. This is not to say that populism is bad but rather is to encourage a more generally educated audience to realize how much of an influence populism can have in politics. I am also optimistic that my analysis of rhetoric in relation to campaigns and movements will provide a deeper level of insight into the power of propaganda.

The Ideology of Populism

Before further exploring the existence of populist campaigns in both Germany and the United States, it is necessary to understand the ideology of populism. In this section, I will define populism and its descriptive characteristics, at times comparing and contrasting populism with nationalism. I will then analyze historic and modern far-right groups in both Germany and the United States, applying these characteristics to the strategies of extremist groups in both countries, then analyzing how rhetoric can contribute to successful populist campaigns. In his book *What Is Populism?* (2016), Jan-Werner Müller seeks to explain populism’s somewhat convoluted existence. He makes the important observation that there is no single “theory of populism” today. Instead, there are general criteria that seek to describe populism.

Leaders of nationalist campaigns serve as symbols and rallying points for the nationalist agenda, focusing more on the culture of a nation and thriving on catchphrases and rhetoric that don’t necessarily rely on the presence of a representative. From this perspective, nationalist leaders serve as the “embodiment of the nation,” whereas populist leaders identify as the “embodiment of the people.” Populism has its roots in popular sovereignty, which is “the contention that the unified will of the people is the supreme authority in a state” (Espejo, 2011, p. 3). Historically, popular sovereignty has allowed marginalized groups to “challenge those in power in the name of a fundamental democratic principle” (Schmidtke, 2023, p. 914).
Today, the way populist politics exercises the notion of popular sovereignty lacks room for the diverse range of viewpoints and people who may exist within that society. Instead, it attempts to identify a singular, united will of the people. Populism often involves a leader who communicates to their supporters that they alone can represent this will. Democratic politics are based in pluralism, which encompasses the “recognition that we need to find fair terms of living together as free, equal, but also irreducibly diverse citizens” (Müller, 2016, p. 8). Populist campaigns will implement antipluralistic rhetoric, which can make way for the warped mindset of the “one authentic people” (p. 9). Populists will then “claim exclusive possession of the people’s political voice; their opponents oppose the people’s voice” (Webber, 2023, p. 855). If the populist figure is the spokesperson for “the people,” then those who oppose the populist figure become the enemies of “the people.” Populism questions the intentions of “pluralistic methods for government decision-making,” as political procedures involving the “expression of diverse points of view” are in direct contradiction with an exclusive will of the people (p. 858). Populists then start to evolve the nationalist-rooted “us vs. them” argument into “moral vs. immoral.”

Oliver Schmidtke states in his article “The ‘Will of the People’: The Populist Challenge to Democracy in the Name of Popular Sovereignty,” that, internally, the will of the people “suggests unity and equality as a promise to its followers,” but externally, it “identifies the enemies against whom decisive action is warranted” (2023, p. 925). By way of repeated rhetorical enforcement, populists will start to ignite a dislike for the perceived enemy, leveraging the nationalist practice of “othering.” In many instances, the “them” in this equation refers to the corrupt, political elite. Populists seek to separate themselves from the corrupt elite who exist above them. Prerna Singh outlines the layers of anti-elitist populism in her publication “Populism, Nationalism, and Nationalist Populism” (2021). The radical right will often combine populism and nationalism to undermine liberal institutions or the cultural elite (Halikiopoulou, 2018). Although both populism and nationalism pose a threat to democracy, the primary difference found between these movements is related to the behavior of the representatives.

Populist leaders will act unorthodox in their methods of speech and rhetoric and push the narrative that they are a representation of “the people.” As opposition parties dismiss populist leaders’ behavior as tasteless, populists can utilize the criticism to their advantage. More recently, with the rise of media platforms as a news source, populist politicians thrive on “criticism from more established candidates and from the media as (it) ... serves to strengthen their followers’ belief in the populist’s authenticity” (Darling & Gatz, 2019). This criticism contributes to the suspicion of and distrust in professional expertise, science, and education (Webber, 2023, p. 863). By labeling adversarial judgment as typical elitist behavior, populists convince their audience that criticism of their nonconformist, unprofessional behavior reflects how these elites view the people: with judgment and distaste.
The very concept of the educated elite can lead to uneducated populations feeling marginalized and results in a pushback against academics and experts. Members of the academic community start to lack credibility, making it much easier for populist politicians to take advantage of political polarization and build upon a distrust of institutions and science. Scientific skepticism is another characteristic of populist politics, and today’s post-truth politics aid in mass-advertising, antiscience messaging. Post-truth politics can be described as a “toxic combination of policy blunders on austerity, war and globalization coupled with a new hybrid media and political system dominated by reality TV, social media, and filter bubbles” (Suiter, 2016, p. 25). As populist politicians embrace a post-truth media world, supporters suffer from a flood of misinformation. Post-truth politics has led to politicians “manipulating popular opinion and discrediting scientific evidence that contradicts their political agendas” (Snodgrass, 2022). Further, using the “us vs. them” narrative, politicians will further emphasize that any ideas that the opposition supports are illegitimate, regardless of their scientific validity.

In addition to naming elites as part of “the people’s” opposition group, populist politics typically erect a group they can perceive as “less than.” This named “underclass” is not actually based on a financial measure of wealth but rather on ascriptive characteristics, including race, religion, sexuality, and immigration status, among others, employed to provide a sense of superiority among supporters. While society’s elites can be hard to distinguish through outward appearance, the characteristics used to identify this group are more readily apparent and provide various identifiable traits toward which supporters can direct their discontent (Singh, 2021, p. 256). This can perpetuate hateful ideologies such as racism, anti-Semitism, and xenophobia and may lead to political violence aimed toward said groups. According to economist Nils Karlson, “The active promotion of political conflict is central to populism” (2024, p. 10), as its existence aids in vilifying the opposing side.

As previously mentioned, populism thrives on simple answers to nuanced, layered issues within society. People want someone to blame during economic hardship. Scapegoating is quite effective in populist politics, as it builds upon the “us vs. them” argument. Typically, the “them” refers to the elites; however, according to Alexander Douglas (2016), the elites “cannot be their scapegoat, since a defining feature of a scapegoat is its inability to retaliate. And the ‘establishment’ is very capable of retaliating.” So, while populist blame may often be aimed at the elites and the establishment, when it comes to action, populist rhetoric often focuses on and leads to attacks on marginalized groups. In her article “The Politics of Fear,” Ruth Wodak states that “all right-wing populist parties instrumentalize some kind of ethnic/religious/linguistic/political minority as a scapegoat for most if not all current woes” (2015, p. 2).
Many populists also take up a strong protectionist stance, which involves the protection of domestic goods, manufacturers, and workers from what is considered “outside” or foreign competition. Politicians who push protectionist views offer up simple solutions to complex economic issues, for example, combining some form of increased public spending with tax cuts (Karlson, 2024, p. 8). Unsustainable economic policy is a byproduct of populist politics’ priority of gaining popularity and emotional resonance among supporters. Often, protectionist and isolationist stances lead to strong anti-immigration rhetoric and policy. Immigration can be viewed by populists and their supporters as the “most intrusive and disruptive (form of globalization) because as a result of it, people are dealing not with objects or abstractions; instead, they come face to-face with other human beings, ones who look, sound and feel different” (Zakaria, 2016, p. 15).

Populism & Nationalism in Germany

The rise of Adolf Hitler and Nazism came at a time when Germany was economically crippled, and experiencing a wounded sense of national pride in the aftermath of World War I. Ideas such as anti-Semitism, expansionism, and Aryan supremacy gained popularity, primarily because of the masterfully effective propaganda techniques employed by Hitler and the Nazi party. In the 1930s, Germany was at its most vulnerable, with widespread suffering as a result of the economic depression, which left about one-third of Germans unemployed (Hall & Ferguson, 2001, p. 132). Further, the damaged national morale in the wake of World War I served as a strategic rallying point for Hitler’s political and military agendas. Hitler also effectively convinced the German public that he would fight for them, metaphorically, by pushing for extreme nationalism, as well as the need for Lebensraum (living space), which he argued was paramount if Germany was going to successfully assume its rightful place as the dominant global force. Hitler and the Nazi party were able to profit from the combination of various misfortunes that plagued German society following World War I.

During Hitler’s tenure as Germany’s leader (Führer), the most influential component of his campaign was his disturbingly effective propaganda. Hitler shaped public opinion through Nazi propagandists, who “learned how to translate fundamental ideological postulates into a continuous narrative of events, a heavily slanted story of good and evil, easily accessible to mass audiences” (Herf, 2008, p. 17). A major key in Hitler’s dictatorship was his ability to gain control over the press, censoring all information that was released to the German public. The NSDAP, the Reich Ministry for Public Enlightenment and Propaganda, amassed roughly 5.3 million members in 1939 (Herf, 2008, p. 19). Nazi propaganda grew to encompass the press, film, music, and German popular culture. By incorporating simple slogans such as “Die Juden sind Schuld” (“The Jews are to blame”), the propaganda made Hitler’s
message clear and simple, reaching both educated and noneducated alike who felt a collective anger following World War I and were desperately seeking a scapegoat.

As World War II continued, Nazi propaganda failed to address the murder of millions of Jews in Nazi concentration camps. Instead, the propaganda aimed to paint the Allied powers and their sympathizers as a “conspiracy of nonequals” who shared a common goal of achieving international Jewry (Herf, 2008, p. 143). This not only furthered Hitler’s pure Aryan race theory with use of the word “unequal” but also fabricated and exaggerated the Allied powers’ goals to invoke fear in the German population. Hitler’s claim that he was the embodiment of the nation and that Jews were a disease within Germany that must be eradicated is decidedly nationalist in nature. Hitler also used eugenics as a scientific basis for his Nazi propaganda, often citing references from a resolution passed by the International Federation of Eugenic Organizations, which accentuated the importance of eugenics for civilized societies (Kuhl, 2014, p. 27). The Nazis used the science of eugenics to support their ideology of German dominance and superiority in furthering their nationalist cause.

Despite Hitler’s alignment with nationalist ideals, some of his propaganda and rhetorical tactics are populist in nature. The key to Hitler’s verbal propaganda was that he kept his speeches short and impactful, using repetitive phrases. Hitler was able to convince his supporters that the entire world was riddled with corrupt Jews, politicians, and socialists who had all come together in a fight against Germany. He would often refer to the pure Aryan race as Volk (the people), to reinforce the notion that they were the true citizens of the German nation and were therefore tasked with fighting for its prosperity, by whatever means necessary. In his speech at the Münchner Lövenbräu in 1928, Hitler said, “The goal of the National Socialist movement is called: People and Fatherland, our slogan is called: Honor, Freedom, and Bread, and the way it is done is called: Fight” (de Saussure & Schulz, 2005, p. 195). This call to action for the freedom of the perceived Fatherland indicated that the issue the country was plagued with was one that affected all Aryans, requiring they take action to save their nation.

When World War II ended and Germany was divided, the Soviets took control of eastern Germany, imposing communist political ideology. The East was pushed so forcibly into communism that a natural pushback of right-wing and Nazi extremism took hold. There still exists an invisible border between the East and the West, even following the fall of the Berlin Wall. In her article “Why Is Eastern Germany So Far Right?” (2018), Anna Sauerbrey refers to East and West Germans as “unequal siblings… the strong one loves his smaller and uglier brother and accepts that his deviant behavior comes from trauma, but he still looks down on him.” This approach helps explain some of the apparent extremism that has continued to build in eastern Germany over the decades since Germany’s reunification.
In western Germany, fascist sentiment witnessed increased scrutiny and stigmatization. German identity was restructured to follow a more liberal agenda (Tanca, 2017, p. 21). Anyone who proposed right-wing or populist policies was characterized as putting effort toward continuing the Nazi party, which Germany so desperately wanted to cleanse its hands of following the devastation of WWII. The so-called “political elites” were responsible for the process of denazification that occurred during the 1950s. Denazification came to include a total rejection of nationalism, as nationalism became synonymous with “a denial of responsibility for Germany’s disastrous past” (p. 22). The repeated smothering of nationalism unintentionally served as a stimulant for the rise of a populist movement that categorized the political elite as a group who not only misunderstood but also disrespected what it meant to be German. This inadvertently paved the way for populist leaders to push the necessary rhetoric to achieve support.

Between 1990 and 2000, right-wing violence in Germany was responsible for more than 200 deaths, with multiple murder and attack reports filed into the early 2000s (Biess, 2020, p. 359). In the year 2016, the BBC reported an average of 10 attacks per day on migrants, with 560 people injured in total (BBC, 2017). In addition to xenophobic hate crimes, Germany has seen anti-Semitic crimes, including the murder of two people outside a synagogue by a right-wing extremist in the east in 2019. Frank-Walter Steinmeier, in an interview with Malcolm Brabant, stated that the neo-Nazis in Germany “want... an authoritative state that aggressively excludes parts of society. They see themselves as part of a tradition” (Brabant, 2020).

Additionally, there are economic factors that can assist in explaining the rise of populism in Germany. The shift in the labor force seen following WWII due to a high influx of migrants and refugees has been directly correlated with the increase in unemployment among German nationals (Tanca, 2017, p. 12). Populist parties have been able to gain leverage among supporters through antiglobalization rhetoric and policy proposals. Populist movements thrive during economic crises. This can explain why there was an escalation in German populism in the wake of WWII, when the country was economically crippled. In recent decades, the rise of populism in Germany has evolved to encompass nativism. Nativist ideology follows “extreme cultural or economic protection of ‘homogenous’ nationals in their nation states against immigrants and the effects of globalization” (Tanca, 2017, p. 8). Nativism attempts to rally supporters against the elite, linking directly to the rhetoric deployed by influential populist figures. Today, this protection of the homeland from the impact of globalization is, in Germany’s case, known as Euroscepticism, involving the strong opposition to the expansion of the European Union and the power that comes with said expansion (Brack & Startin, 2015, p. 1).

The Alternative für Deutschland (AfD), a far-right wing political party, is an example of a modern form of populism in Germany. The AfD first appeared in 2013 but
grew in prominence in 2017, when it was able to secure 12.6% of the vote, as well as 92 seats in the Bundestag (Gedmin, 2020). The AfD holds strong anti-immigration and antirefugee sentiments. For some time, the party was able to convincingly pass as a moderate group; however, because of an increase in racist nationalism and anti-Semitism, the AfD has shifted into the alt-right category. The party finds its most ample supporters in eastern Germany and includes both the “right-wing extremist fringe and people dissatisfied with the status quo” (Chase & Goldenburg, 2019). The rise of this party is consistent with some of the global trends surrounding populism that have also manifested in the United States. Leaders of the AfD routinely dismiss the mainstream media as complicit with the corrupt left-wing political elite. The AfD overlaps populism and nationalism in its messages, which challenge immigration as an onslaught of foreign agendas that will weaken the nation and its values.

The largest groups in Germany that experience backlash from the AfD are Muslims and immigrants. A popular form of rhetoric for the AfD, as well as other far-right extremist movements, is known as femonalist rhetoric, in which there is a “distortion of feminist ideas … for the stigmatization of migrants, ethnic minorities, and Muslim women through racialized representations of these groups as inferior” (Doerr, 2021, p. 4). For example, a poster released during the AfD’s 2017 campaign showcased two women in bikinis. The message said, “Burkas? Wir steh’n auf Bikinis” (“Burkas? We like bikinis.”). This is a blatant criticism of Islamic tradition, posed as an attempt to appear sympathetic to the feminist movement. This rhetoric attempts to use the assumed oppression of Muslim women who wear Burkas as a flaw in Islamic tradition in relation to the modern feminist advertisement of the “empowered woman” who has control over her body (p. 4).

Recently, the AfD’s cochair and spokesperson, Joëg Meuthen, has stepped down from his role, claiming he can no longer support the “clear totalitarian echoes” that the party has come to represent (Deutsche Welle, 2022). Meuthen has spoken out against the extremist tendencies that the AfD has gravitated toward, and he has been met with opposition from various members of the party. He also lists the AfD’s cult-like approach to the COVID-19 pandemic as a driver for his resignation (Deutsche Welle, 2022). The AfD has incorporated Islamophobia into its propaganda, criticizing traditional practices of Islam and mixing in left-wing rhetoric, in this case rhetoric associated with feminism, to appear more sensitive to progressive movements.

In addition to offensive messaging within its propaganda, the AfD has been considered morally responsible for various attacks aimed at synagogues and mosques around the country. Many in opposition to the party feel that, through the widespread propaganda and hateful rhetoric, the AfD has been indirectly inciting violence against these scapegoated groups in society. Although the AfD may not be directing these members of society to commit such massacres and atrocious acts, the group is circulating dangerous propaganda that may mark the group as complicit. In 2019, the
leader of the Free Democratic Party in Germany, Mark Buschmann, reported that an attacker at a local synagogue was hunting down scapegoats, and claimed that any political parties utilizing “scapegoat theories” to further their agendas were partially responsible for these beliefs culminating in violence (Witting, 2019). With the rise of the AfD and a growing fear of immigration in Germany, right-wing acts of violence rose by 50% between 2015 and 2016 (Biess, 2020, p. 359). Politicians in Germany recognize the threat that lies within far-right wing rhetoric and hope that the AfD will work to shift its propaganda so it doesn’t put certain groups in society at high risk of acts of domestic terrorism.

Since the 1970s, right-wing populism in Germany has been a movement founded on fear, which has evolved to include fear of Muslims and migrants, as well as governmental/political institutions. The distinction between right-wing populists and conservatives is that populists fear the political elite. In comparison, conservatives tend to fear the fall of authority at the hands of left-wing groups. The two groups can overlap when it comes to fear of the “other,” usually minority groups. Right-wing populist groups “personalize their objects of fear . . . (with) the potential to turn into hatred and acts of violent exclusion” (Biess, 2020, p. 359). According to The Economist’s 2021 election statistics report, the AfD party was able to obtain 83 seats, or 10.3% of the vote. Although this was a lesser share of the vote than in 2017, AfD has a concerning presence in the German parliament that should not be underestimated.

Populism and The United States: A Historical Overview

Populism has existed in American politics since the mid-19th century. The Know Nothing party is one of the earliest recognized nativist/populist political parties in the United States. The party found initial success following the mass Irish immigration in wake of the potato famine (Alsan et al., 2020, p. 8). Led by Thomas Whitney and William Poole, the party embraced both nationalism and religious discrimination (Boissoneault, 2017). One of the Know Nothing’s primary goals was “reducing the immigrant threat to native workers” (Alsan et al., 2020, p. 10). Fears surrounding job stability because of the influx of immigrants and industrialization, as well as anti-Catholic sentiments, contributed to the party’s influence (p. 9). Eventually, the party collapsed because of its inability to take a position on abolition, but the party’s approach to immigration “has been apparent in policies aimed at each new wave of immigrants” (Boissoneault, 2017).

In the late 19th century, the Populist Party was formed. This party was led by Thomas E. Watson and sought to meet the demands of the southern-based Farmers’ Alliance, whose primary belief was that the political economy of the United States served only the rich (Shaw, 2020). The Farmers’ Alliance called for the government to implement inflationary policy, in hopes of increasing the price of cotton (Shaw, 2020). The party sought the support of African Americans in its political campaigns. The
presidential election of 1896 did not prove successful, and eventually, the party collapsed. Thomas Watson, however, returned to politics in the early 20th century and advocated for depriving African Americans of the right to vote in Georgia. This led many African American voters to feel that “the movement offered them little and Populist appeal had more to do with opportunism than friendship” (Shaw, 2020). The fallout from Thomas Watson’s betrayal of the African American voter was influential in populism’s relationship with race.

In the 20th century, Father Charles Coughlin’s radio show became “the voice of the people” in America following the Great Depression (Wang, 2021, p. 3065). Radio was a relatively new form of broadcasting and introduced a new medium with which to spread influence. Father Coughlin’s history as a religious speaker lent itself to his success as a compelling radio host. In the 1930s, he became “a leading anti-Semitic, icon, fascist sympathizer, and isolationism advocate in prewar America” (p. 3086). According to Wang, Coughlin can be considered the first populist media personality in the United States. He had a persuasion rate of 28% among the American public and influenced the formation of pocketed groups of pro-Nazis (p. 3089). Prior to radio, populism had existed primarily as a political ideology. Father Coughlin demonstrated how the adoption of new technology media would forever change the ability of populist politicians to garner support.

Populism has continued to emerge in American politics in more recent decades leading up to the 2016 presidential election and has found strong footing following economic crises. The right-wing Tea Party movement gained popularity following the 2008 financial crisis, and in 2010 had achieved support from almost a quarter of the American public (Minkenberg, 2011, p. 290). The party was majority White and middle-class and promoted antiestablishment messaging. Many members of the movement felt strongly that the political elite fashioned projects “aimed at concentrating in the hands of a ‘small group’ of putative experts ‘an almost complete control over other people’s money, other people’s labor, other people’s lives’” (Rahe, 2011). Some have identified the Tea Party as fostering the prevention of a practical solution to the debt crisis in 2011 (Minkenberg, 2011, p. 283). After the 2011 financial crisis, the left-wing activist movement Occupy Wall Street gained traction, focusing on issues such as the climate crisis, growing wealth disparities, debt, and rising housing and healthcare costs (Levitin, 2021). The group’s motto was “We are the 99 percent,” pointing a finger at the wealthiest 1% and shedding light on the reality of the wealth gap in the United States. In the end, Occupy Wall Street was unable to prevail in the face of poor weather, weariness, and tension among occupiers (Anthony, 2021). Even following these movements, productivity growth was slow, and many Americans felt that “many of the emergency measures adopted in the aftermath of the crisis . . . benefited the well-connected few at the expense of everybody else” (Rohac, 2018). American populism surrounds the “mistrust of major institutions . . . and suspicion of
global elites” (Rohac, 2018). Those who are loyal to populist sentiments in the United States associate the Great Recession with the wrongdoings of the elite.

In the United States, democracy’s strength has become questionable as people continue to distrust government institutions. Additionally, the structural changes seen in the labor market during the early and mid-2010s has resulted in outsourcing of labor to low-cost locations. The ensuing threat to economic stability, as well as the American value placed on individualism and the dignity associated with maintaining a stable job, contribute to populism’s success. Like the institutionalized and enforced embarrassment associated with nationalism in Germany following the war, those in the United States who can’t seem to find themselves in a stable economic position feel humiliated by the entity that is supposed to look after them: the government. This “dignity deficit” is “a potent resource ripe for unscrupulous political candidates to translate into popularized anger” (Rohac, 2018).

The loss of status further exacerbates the gap between the corrupt, political elite and the ordinary people. In the United States, populists often criticize the political elite for catering to the needs of minorities while overlooking their suffering, which provides the resentment of “the other” in American society. In the United States, while immigrants have been at the center of hateful rhetoric, it is African Americans who face the most discrimination as related to the populism movement (Rohac, 2018, para. 15). African Americans still suffer the long-term ramifications of slavery and segregation, as seen in the systematic inequality that exists in economic, social, and health institutions today. Populists “exploit anxieties related to… demographic change” (Rohac, 2018, para. 16). By building upon the fears of the people, American populist politicians can easily provide scapegoats for discrimination.

Leading up to Donald Trump’s election, many components, including economic insecurity, terrorism, globalization, and various other influences, promoted the 2016 trends in the American electorate. James McGann, a senior lecturer at the Lauder Institute University of Pennsylvania, names economic, physical, and information insecurity as contributing factors, as well as a loss of national identity and a lack of confidence in government institutions. Globalization plays an important role in the loss of national identity and served as a foundation for Trump’s campaign slogan “Make America Great Again,” with the focus on “America First,” rallying to his side voters who desperately felt the need for a resurgence in American patriotism. These factors, when combined with various other ills and perceived oppression by the American conservative party, “drive populism and politicians who tweet simple solutions to complex problems” (McGann, 2016). Hitler also utilized the plainest of slogans to unify the German people under a simplified, unrealistic ideology.

Both Donald Trump and Bernie Sanders called for pushback against the establishment. Anti-elitism was present in both campaigns; however, Bernie Sanders’s populist agenda had a slightly different underpinning, including the general
questioning and suspicion of the extent to which the current institutions truly served the ordinary American people. Donald Trump, while supporting anti-elitism despite being a member of the elite himself, took the argument a different route by repeatedly reiterating an unconventional candidacy that reflected the people of America, who had been “misunderstood” for far too long. Both candidates employed populism in their campaigns, yet only one succeeded.

**Populism: Donald Trump’s Campaign and Presidency**

Pauline Jones and Anil Menon of the University of Michigan developed criteria to classify different types of populists (Jones et al., 2019). Jones and Menon use two major criteria to classify these figures. First is the figure’s position in the political landscape as either an insider or an outsider. The second measure follows the individual’s ideological commitment, and whether the individual subscribes to opportunism or true belief. Leading up to Donald Trump’s campaign, he was both a political outsider and an opportunist. According to Jones and Menon’s chart of classifications, Trump is a strategic populist, targeting his rhetoric solely at the masses and failing to align his rhetoric with effective policy. Leading up to the 2016 election, the influence of the economic crisis on the lower-income, blue-collar areas of the country was impactful enough to allow for strategic populist rhetoric to gain a foothold among the population. During his campaign and presidency, Donald Trump consistently used a combination of the arguments that the United States, once special, is now a weak player on the world stage and that this humiliation has been brought about by America’s enemies, including both internal and external figures. Trump would often “arouse a sense of threat and then resolve it by providing an answer in the form of himself” (Rowland, 2021, p. 25). He successfully convinced supporters that he is “representative of the group in both a symbolic and a practical way, able to represent the group at the political level” (Reicher & Haslam, 2017). Trump was able to tap into the collective unconscious surrounding the “us” and falsely reinforced his seat in the room among the “alienated” members of American society.

Trump lives an outwardly rich lifestyle, one that the masses associate with the attainment of the American dream. Donald Trump is nowhere near what society would consider the struggling working-class American; therefore, to appear more legitimate as the voice of the common people, Trump addressed his wealth in a way that proved advantageous, arguing that he “has been so successful and become so rich that he cannot be bought” (Reicher & Haslam, 2017). This self-characterization of his wealth juxtaposes the various other elites who are easily corrupted by financial incentives. Trump claimed to “tell it as it is,” presenting himself as a nonconformist in contrast to the existing convoluted political landscape filled with insincere, half-promises from the more established candidates. Most of Trump’s charisma came from the fact that he almost always seemed to be unprepared in his speeches, which gave
him an air of approachability (Khazan, 2016). In this way, Trump continued to exercise
the argument that he is the embodiment of the “true” American people.

Trump “creates a sort of authoritarian godlike aura” as the one who will
“protect all of your rights and all of these huge issues around justice, fairness and
freedom of speech” (Khazan, 2016). He curated the “us vs. them” mentality through his
speeches, and the message behind “Make America Great Again” is synonymous with
the return to a “golden age,” in which the real, moral citizens of the nation triumph. By
painting his supporters as the morally righteous members of society, Trump was able
to make the claim that everyone else was not just wrong but morally evil. In right-wing
populism, insult politics aid in labeling the moral vs. immoral using “norm-breaking
language (as a) political strategy” (Winberg, 2017, p. 3). Trump often employed insult
politics by using nicknames such as “Crooked Hillary,” “Cheatin’ Obama,” and “Lyin’
Ted” (Relman, 2019). By attacking his opponents aggressively, he was more likely to
receive pushback from them. Trump was then able to use their disapproval of him as
the unconventional leader as further proof that the political elite are snobby and
judgmental. In an interview with Ari Shapiro of NPR, Jennifer Mercieca, an American
political rhetoric historian, points out the concern when dehumanizing wordplay is
utilized in front of mass audiences. Mercieca states that “the only time you see
presidents using the rhetorical strategy… (of) treating people as objects is when
they’re using war rhetoric” (Shapiro, 2021).

Much of Trump’s rhetoric played on the fear that his followers have
surrounding a variety of issues, including the implementation of a “socialist agenda”
by the liberal elite, as well as a loss of a core American identity and values. Trump
supporters were convinced that if an opposition group won, they would lose their
freedom. Trump exaggerated the reality of the “liberal agenda” to bring fear to his base,
so much so that many supporters were convinced that being White, Christian, or male
made them oppressed or discriminated against by other groups. Donald Trump also
advocated for the protection of American jobs during his campaign, and he enacted
protectionist policies during his presidency. As previously mentioned, populism often
results in protectionist and isolationist stances. To further contribute to his
“Protector of the People” role, Trump claimed that America must “protect our borders
from the ravages of other countries making our products, stealing our companies, and
destroying our jobs” (Irwin, 2017, p. 45). This rhetoric evolved into the “America First”
trade policy, which Trump claimed would lead to prosperity and strength (Irwin 2017,
p. 45). By enacting an “America First” trade policy, Trump dispelled the economic
concerns of his followers by reinforcing himself as their formidable hero and
protector; however, U.S. trade policy expert Douglas Irwin states that the “America
First” trade policy would “do nothing to create new manufacturing jobs or narrow the
trade deficit,” and instead risks “triggering a global trade war that would prove
detrimental to all countries” (2017, p. 45). Further, a study by the U.S.-China Business
Council (USCBC) estimates that Trump’s trade policies resulted in a net loss of almost

99
245,000 American jobs (Pettis, 2021). Note that this outcome is consistent with the aforementioned unsustainable economic policy that often accompanies populist leadership.

Donald Trump also repeatedly made efforts to “delegitimize core democratic institutions” (Savelsberg, 2020). Trump utilized populist rhetoric to stir a strategic distrust among his supporters in the democratic processes of the United States. By expanding the distrust of elites to encompass governmental institutions, he effectively used the already existing “distrust in the political system as a political weapon” (Fried & Harris, 2020, p. 528). Trump routinely referred to the FBI’s search of his Mar-a-Lago home as a “witch hunt.” The use of this rhetoric resulted in death threats and attacks on FBI officials in Cincinnati and Florida (Stone, 2022). As mentioned in my earlier literature review of populism, promoting political conflict is central in populist success. Trump implemented the argument that an illegal voter-fraud scheme involving mailbox robberies, forged ballets, and fraudulent voter signatures by the left was to blame for his loss in the 2020 election (Fried & Harris, 2020, p. 527). Distrust in the political system, paired with the belief that the opposition actively took measures to prevent the votes of conservative voters from being counted, creates an angry base susceptible to even more extreme future populist rhetoric and propaganda.

Donald Trump’s Use of the Scapegoat

A key characteristic of populism is the tendency of the spokesperson to identify a scapegoat. Because of Donald Trump’s routine scapegoating of immigrants, reluctance to condemn White supremacists, and insistence upon calling the coronavirus some variation of “the Chinese virus,” Donald Trump not only left specific groups of society in a dangerous position but has additionally mobilized and paved the path for far-right hate groups that have proven they are ready to commit acts of violence. Trump used the scapegoat theory to further his anti-immigrant rhetoric. This approach attempts to provide someone to blame for the “underserved Americans” who feel cheated out of jobs. By condemning immigrants and other marginalized groups as the foundation of America’s ills, Trump united the extremists looking for the “other” toward which to direct their hatred. Othering, when paired with anger, has the dangerous likelihood of culminating in violence. Within his speeches, Donald Trump used extreme language to convey simple messages, often exaggerating the situation to resonate with the anger his supporters felt. Instead of addressing immigration in a policy-oriented way, Donald Trump, during his State of the Union speech in 2019, claimed that “countless Americans are murdered (year after year) by criminal illegal aliens” (Lach, 2019). This statement is false yet is effective at garnering a strong, emotionally charged response from his supporters. In addition to referring to them as criminals, Donald Trump has called undocumented immigrants “rapists and gang members who pose a direct threat to the welfare of ‘law abiding’
people” (Perea, 2020, p. 1). By using words such as “invasion,” Trump implies that immigrants are an enemy that is infiltrating America, taking American jobs, and “freeloading.” As a result, Hispanic Americans became a targeted group. Anti-immigrant rhetoric aimed at the southern border also appeals to white supremacists’ “prevailing conception of the United States as a country controlled and dominated by whites and their culture” (Perea, 2020, p. 2). The combination of populist anti-immigrant appeal and white nativist beliefs poses a threat to the safety of Hispanic individuals living in the United States. In 2019, the FBI reported a 21% increase in anti-Latino and anti-Hispanic hate crimes (Canizales and Vallejo, 2021, p. 155). One of the most extreme anti-Hispanic hate crimes in recent years took place in 2019, when a white nationalist murdered 22 people at a Walmart in El Paso, Texas. El Paso borders Mexico, and its population is roughly 80% Latino (Canizales and Vallejo, 2021, p. 155). The claim that illegal immigration is a crisis and that Hispanic citizens are rapists and gang members is inaccurate. In contrast, undocumented immigrants are “statistically less likely to commit crimes than native-born Americans” (Perea, 2020, p. 2). Further, immigrants residing in the United States contribute billions in taxes each year and often take “undesirable, low wage jobs” (Perea, 2020, p. 2).

The United States has also seen a rise in hostility toward African Americans in recent years. Trump’s use of the collective “us vs. them” argument, as mentioned in the earlier populism review, results in members of the “us” looking for physical characteristics to associate with the “them.” In the wake of the murder of George Floyd by Minneapolis police officer Derek Chauvin in May of 2020, Black Lives Matter (BLM) protests erupted not only in the United States but around the world. In response to the protests, members of Trump’s base pushed back with misplaced hostility, claiming an anti-protest rhetoric, with the argument that violence and riots should not be tolerated and would “desecrate” the memory of George Floyd. Others took things a step further; in Kenosha, Wisconsin, 17-year-old Kyle Rittenhouse shot and killed two protestors and injured another (Bekiempis & Gabbatt, 2020).

Trump’s use of scapegoating usually accompanied a shortcoming in his own leadership and policy. The effectiveness of division politics proved to work during his campaign, so it only makes sense that he would turn to these tactics during his presidency. During the height of the COVID-19 pandemic, Asian Americans became the victims of Trump’s “othering.” Common terms used to refer to the virus were “Chinese virus,” “Wuhan virus,” and “Kung Flu” (Porumbescu et al., 2022, p. 2). Trump’s insistence upon using these terms can be considered partially responsible for Asian hate crimes. Between the time the first case of COVID-19 was recorded in the United States and 2021, there were 3,800 anti-Asian hate crimes, as reported by NBC (Yam, 2021). Trump’s public criticism and verbal attacks aimed at China and its government’s response to the pandemic allowed for anger to be “redirected toward ethnic groups viewed as representative of that government” (Porumbescu et al., 2022, p. 2).
Donald Trump also incorporated islamophobia into his scapegoating, using Islamic extremism as a basis for tighter national security. In discussions about Islam, Trump “compared the religion to a ‘malignant cancer’ and tweeted that a fear of Muslims is ‘rational’” (Zurcher, 2017). Utilizing dehumanizing rhetoric in describing the religion of Islam, Trump painted Muslim individuals as “less than human” and therefore not in possession of human emotions or worthy of human sympathy. Additionally, Trump’s call for a Muslim travel ban led to fear among American Muslims, with members of the faith “afraid to wear headscarves . . . to speak Persian in public . . . for fear that they will be targeted as terrorists” (Katirai & Puig, 2017). Similar to other victims of Trump’s scapegoating, Muslim people experienced an increase in hate crimes leading up to and during his presidency. It was reported that, following Trump’s election, between January and March of 2017, there was an average of one anti-Muslim incident every other day (Patel & Levinson-Waldman, 2017).

I propose that there is a slight disconnect between the argument Trump is making and what his followers feel they are supporting, however. Trump is a populist who is acting sympathetic to the nationalist cause. In a way, he is enacting a nationalist-populism rhetoric. Many of his followers are nationalists yet can’t differentiate between nationalism and populism when it comes to Trump’s promises. A significant number of these far-right wing extremists aren’t even aware of the ideology of populism and can’t objectively look at Trump’s rhetoric in a populist light. Instead, supporters are willing to take anything Donald Trump says and turn it into permission to act in what they believe is the nation’s best interest, even if it involves attacking the United States’ fundamental institution of democracy. Whether or not Trump outwardly acknowledges the correlation between his populist rhetoric and the nationalist-motivated violence that increased in America during his presidential term, his rhetoric was partially, if not fully, responsible for various domestic terrorist attacks against scapegoated groups. As previously mentioned regarding violence in Germany with the AfD’s emergence and growing popularity, using scapegoat theories sooner or later results in a sense of duty in supporters of the right-wing extremist movement, who feel a “call to action.”

The major element making Donald Trump’s campaign populist in nature is that many of his claims and arguments are not backed by science. Populism, as explained previously, is an ideology used by candidates who want to appeal foremost “to the people” rather than to any politically or scientifically accurate cause. I therefore argue that the pitfall of populism in terms of legitimacy is its inability to be backed by any scientific evidence. Donald Trump’s campaign was fueled solely by opinions and the backing of his supporters. There was no scientific support for any of Trump’s rhetoric or arguments; therefore, he can only be populist. Trump’s campaign surrounded his popularity as a candidate and fighter for the people, which doesn’t need or warrant factual or scientific support. Nationalist candidates, in contrast, run a different type of campaign—one that doesn’t focus so much on the popularity of the leader as on the
popularity of national pride. There is therefore more room to introduce scientific arguments into a nationalist campaign—for example, Hitler using eugenics as a “scientific” basis for the crimes committed against European Jews. This is what differentiates Donald Trump’s populist rhetoric and propaganda from Hitler’s more nationalist advocacy and is also what differentiates the two as political figures.

Conclusion

All of this then raises the question of whether it is necessary to have a cataclysmic event with extremists at the forefront to result in the rejection of them as a group—and, furthermore, just how cataclysmic would this event have to be? Although I conclude my study with a focus on the United States and the ongoing influence of the alt-right, I have mentioned throughout that the alt-right has continued to grow in popularity, support, and political influence in Germany. Given Germany’s history, it is surprising to see the AfD continuing to gain support. The tensions resulting in the rise of the extreme right in Germany are not an isolated European event. In April of 2022, France saw a presidential election with Marine Le Pen trailing behind Emmanuel Marcon by only around a 5% marginal difference. Le Pen is a far-right leader who espouses anti-immigrant and nationalist ideologies. When one considers the possibility of a cataclysmic event, it thus could well occur on a global scale.

An inherent power is found in Trump’s supporters, lying in the fact that many of the members of his base are collectively united through hatred, or anger. This is attributed to an array of reasons, including, as highlighted in this work, the propaganda and rhetoric Trump has employed and his ability to appeal to emotions and to provide seemingly simple solutions to the various complex ills that plague the people who make up his base. This unity is what makes the extremist groups on the right almost impenetrable in their views. One could argue that anger has also been at the base of some of the most effective activist movements by left-leaning parties throughout history, most recently seen with BLM. The outrage following the murder of George Floyd by police officer Derek Chauvin served as a catalyst for one of the largest antiracist movements since the civil rights era of the 1950s and ’60s.

The major difference between the right and left in the United States is the mere scope of issues that the left encompasses, which makes it difficult for various subgroups to unite on one common goal. There are activists and supporters on the left who support a whole range of issues. This includes those who stand up for LGBTQ+ rights, those who fight back against discrimination whether based on race or gender, climate change activists, and those who support anticapitalist sentiments. The left is made up of many different circles of ideologies, which can sometimes pose a problem when it comes to uniting a base with one common objective. The familiar discontent felt among Trump supporters allows for a coalition that results in loyalty and
admiration, which may be why Trump supporters have appeared more prone to buying into the Trump base through gear including flags, hats, and stickers. On the other side, many left-leaning voters don’t feel a significant attachment to Biden, instead claiming he received their vote solely to achieve Trump’s removal from office. A poll conducted by Pew found that 56% of Biden supporters voted for him because he was not Trump (Brewster, 2020). The lack of ability for Biden voters to pinpoint specific reasons why they voted for him outside of their distaste for Trump is what shows a slight weakness in the ability of the left to mobilize with one purpose. The BLM protests were a recent example of how impactful such a goal could be. Overall, the left is more dispersed in terms of goals, and this results in a disadvantage when put up against the collective subconscious of the right.

Donald Trump’s rhetoric laid the proper groundwork for right-wing extremists in the United States to continue without his leadership in pursuing their ambitions. Trump was effectively able to cast doubt upon any outsiders, on both ends of the spectrum. On the one side, he was able to convince his supporters that the corrupt liberal elite will always look down on them and that any questioning by leftist politicians of members of the right stems from an inherent disrespect. On the other side, Trump gave his followers groups of people to scapegoat and blame for their ills. The populist appeal and rhetoric outlined earlier in this work transcend Donald Trump as a leader and will continue to influence the right-wing extremist movement, with or without him in power.

A combination of factors leads to my conclusive thoughts regarding a cataclysmic event initiated by the right. When Donald Trump was banned from Twitter in 2021, his supporters were outraged. Although Twitter has the right to suspend or ban any of its users if it deems the users’ content inappropriate, many felt that this decision was an infringement upon the First Amendment regarding freedom of speech. Despite good intentions by Twitter and a reasonable cause for the permanent suspension of Donald Trump’s account, this action may have indirectly fueled the extremist fire. No matter what comments were being silenced, even if they did enable or incite violence, supporters viewed this as a violation of free speech. With their leader’s free speech being inhibited, and him unable to exploit his propaganda, many right-wing extremists questioned the legitimacy of democracy, which gave rise to the eerie concept of the “silent majority.”

The concept of the silent majority communicates a presence that exists but whose influence has not yet been felt. When Trump supporters claim to be members of the silent majority, the claim is usually accompanied by a threatening undertone. The reality is that this group of people is unlikely to constitute a majority. The notion of the silent majority isn’t necessarily literal but rather serves as a force that empowers members of the extremist right. With the backing of the invisible silent majority, regardless of its legitimacy, right-wing extremist members may feel that if
they act in accord with their beliefs, they will have support. This idea is also very menacing in nature because it communicates a larger force than may actually be existent to those who would oppose members of the right. There is something elusive about the whole concept that incites fear in opposition groups. Now, the silent majority may require more serious consideration, considering the Capitol riots, where we witnessed an assault on democracy following the 2020 election. Additionally, the testimony given by former White House aide Cassidy Hutchinson in June of 2022 furthers the notion that Donald Trump’s rhetorical tactics are intended to incite violence and that the former president was willing to turn a blind eye while his supporters stormed the Capitol. Furthermore, Donald Trump’s indictment in March may lead to unanticipated reignited waves of support for him and could even prove fruitful for him gaining supporter sympathy.

One of the more recent concerning developments that may contribute to a catastrophic outcome is the anger surrounding the 2020 election. The pushback against Biden’s legitimacy as the president-elect was introduced in the days and months leading up the election night when Donald Trump voiced strong opposition to mail-in voting. Mail-in voting has always been used in presidential elections and became even more prevalent with the COVID-19 pandemic, as voters did not necessarily feel comfortable flocking to the polls during a period of such high infection numbers. Despite no significant proof of mail-in voting fraud, Donald Trump routinely made false claims surrounding the practice. He spread misinformation such as the notion that counting ballots after election day is evidence of a “rigged system,” or that fraudulent unsolicited ballots were being counted (Parks & Karson, 2020). In addition to these statements, Trump continued to instill distrust in the system by alleging that military ballots with Trump’s name on them had been discarded in wastebaskets in Pennsylvania. A few hours into the election, there was a sharp increase in blue votes, since mail-in ballots are counted later than in-person ballots and Democrats are more likely than Republicans to vote by mail (Coleman, 2020). As the tide started to, seemingly overnight, turn toward a Biden victory, Donald Trump had already laid the groundwork for delegitimizing any outcome that did not result in his victory.

The reason this is so important in analyzing and predicting what is to come from the extremist right is how it works in conjunction with the various other misfortunes of which Trump supporters feel they are victims. The anger that possesses this group, and the need for a scapegoat, already makes them a threatening force. These are people who feel cheated out of success and who feel ignored by the liberal elite in favor of minorities, despite feeling that they themselves are the real, hardworking backbone of America. This perceived disservice is problematic on its own, and the “silencing” of the leader who stands up for them, the person who sympathizes with them on a national stage and makes them feel heard, fuels the rage that was already present. And finally, the notion of the election being “stolen,” which has led to a lack of faith in democracy as an institution and eventually led to the attack
on the Capitol, showcases just how violent a group like this has the potential to become. No matter how horrifying the attack on the Capitol appeared, the silent majority, existing quietly among us with these collective feelings of outrage, has potential to be extremely dangerous for opposition groups and democracy.
Bibliography

Anthony, A. (2021, September 12). “‘We showed it was possible to create a movement from almost nothing’: Occupy Wall Street 10 years on. The Guardian. www.theguardian.com/us-news/2021/sep/12/occupy-wall-street-10-years-on


Brabant, M. (2020, November 18). How right-wing extremists have infiltrated German security forces. PBS. www.pbs.org/newshour/show/how-right-wing-extremists-have-infiltrated-german-security-forces


Friedman, G. (2016, March 10). 9 times Donald Trump was compared to Hitler. *The Times of Israel*. www.timesofisrael.com/9-times-donald-trump-was-compared-to-hitler/


McGann, J. G. (2016). Why Donald Trump won the election and does it mean the end to think tanks and policy advice as we know it? Penn Collection. https://repository.upenn.edu/ttcsp_papers/6


Yam, K. (2021, March 19). *There were 3,800 anti-Asian racist incidents, mostly against women, in past year.* NBC News. www.nbcnews.com/news/asian-america/there-were-3-800-anti-asian-racist-incidents-mostly-against-n1261257


STREAM GAUGING: INVESTIGATING THE FLOW RATE AND RESIDENCE TIME IN LAGUNA BACALAR, MEXICO

ELIZABETH J. KOWALCZYK, MIDDLE TENNESSEE STATE UNIVERSITY
MENTOR: TIMOTHY GRUNDL

Abstract

Stream gauging is a standard tool used to measure the flow rate of various bodies of water. Knowing the flow rate of a body of water allows for the residence time of a body of water to be calculated. Currently, there is very little information regarding the flow rate and residence times of Laguna Bacalar. This system has been observed hydrologically consistently only in the past 6 years. In this study, 13 stream-gauging locations of interest were identified and gauged for their respective flow rates using the midsection method. Once the flow rates were acquired, maps depicting the Laguna Bacalar flow rate for November 2022, summer 2023, and the overall average were created. Following this, the residence time for the main four basins found in Laguna Bacalar was calculated. Understanding the overall system is vital to understanding the hydrological ecosystem, which depends on these factors.

Introduction

Lagunas are bodies of water found all around the world. They come in all shapes and sizes and play a massive role in biodiversity and ecological sustainability. Lagunas are bodies of water where river and ocean life meet, separated by a natural barrier (Li et al., 2020). They serve as homes for many unique and diverse animal, fish, and plant populations (Li et al., 2020). Typically, lagunas are calm and shallow, providing a perfect location for various organisms. Many fish populations use laguna environments as breeding grounds and early habitats for their young (Li et al., 2020). Lagunas also serve as natural filtration systems. They trap sediments and pollutants, preventing them from reaching open waters, where they could do more damage (Davies-Vollum & Boateng, 2023). The importance of lagunas lies in their role as dynamic ecosystems that support a variety of ecological functions while also offering benefits to human societies (Davies-Vollum & Boateng, 2023). Conservation and sustainable management practices are crucial to ensuring the continued health and resilience of these valuable environments.
The average length of time that water remains within an ecological waterway, known as the hydraulic residence time, is a key parameter controlling a system's behavior (Gilboa et al., 2022). The hydraulic residence time is defined as the volume of water in a reservoir divided by the rate of addition to or loss from it (Brittanica, 2019).

Residence time affects multiple complex processes such as carbonate distribution and influence in a marine environment (Aquilina et al., 2003), soil's nutrient distribution within a water source (Ma et al., 2019), and overall water quality within a system and how that affects people (Meals et al., 2009), and much more. The biogeochemical behavior of waterways and other aquatic environments often depends on the physical processes of transport and mixing (Rueda et al., 2006). The transport and mixing that occur help determine the spatial location, distribution, and fate of dissolved natural and suspected substances such as pollutants, along with indicating important environmental conditions (Rueda et al., 2006). Residence time estimates can be compared with time scales of inputs or biogeochemical processes to understand populations and chemical properties of a given waterway (Rueda et al., 2006). Many lakes, especially in remote or semiremote locations, are monitored either very little or not at all (Bruckner, 2017).

Stream gauging is a practical tool to measure waterflow moving through a given stream per unit time (Bruckner, 2017). This not only helps with understanding residence time but also helps with maintaining multipurpose water management systems, issuing flood warnings, mapping floodplains, and protecting water quality (Bruckner, 2017). Stream gauging was used to understand the stream flow and residence time of Laguna Bacalar.

Laguna Bacalar is the second largest laguna in Mexico and the largest lagoon in the Yucatan Peninsula. Renowned for its colors, Laguna Bacalar is often referred to as the Lake of the Seven Colors and stretches for more than 42 km. Laguna Bacalar is being subjected to an increase in tourism with the discovery of its unique freshwater microbialite reef (Elsworthy, 2022). Since living bacterial mats, called stromatolites, became popularized in the 2000s, people from all over the world have come to view the feature (Elsworthy, 2022). In 2015, Mexico's federal environmental protection agency issued a pollution alert for the laguna, citing human contamination concerns (Nast, 2021). Concerns over the connection between the pollution and the health of the microbialite reef have arisen.

In addition to being home to many unique freshwater features, such as stromatolites, the lake also serves a wide range of purposes. Culturally, the Mayan people hold the lake in high regard. They have been using the lake for thousands of years for fresh water, for food in the form of fish, and as a spiritual location (Culturetrip, 2017). Today, the lagoon is still used for these purposes in addition to being a tourist destination to increase the economy of the small town (Culturetrip, 2017). Locals have pointed out the possible connection to the recent “browning”
discoloration of the lake with the increase in tourism (Nast, 2021). With more tourism, protecting the laguna has become a vital priority of the community.

Given that the concerns with the lake’s health have arisen only recently (around 2015), Laguna Bacalar’s flow rates have been measured by our team only over the past few years, starting in 2017. This appears to be the first time such a study has been conducted at this location. Understanding Laguna Bacalar’s residence time leads to a better understanding of the overall health of the water. This will help researchers better understand how quickly the laguna will be able to recover from the impacts of increased tourism and can help with additional research projects that may focus on the carbonate or pollutant concentrations, by offering a diagram that depicts the water flow. By combining and evaluating years’ worth of data, starting in 2017, we aim to explore this research question by conducting stream gauge measurements to quantify stream flow and water residence time of Laguna Bacalar. With this new information, the government will have a better understanding of the lake’s health and have a better idea of how to approach protecting the laguna without harming the economic benefits the laguna provides. Figure 1 depicts Laguna Bacalar and its location within the transverse coastal corridor, which is a complex ecosystem defined by its interconnectivity of hydrological, biogeochemical, and ecological interactions (Hernández-Arana et al., 2015).

Figure 1. Map of Laguna Bacalar Within the Transverse Coastal Corridor
Materials and Methods

Steam gauging of Laguna Bacalar first started in 2017 with two randomly selected locations within the laguna. In 2022, an additional seven sites were chosen based on several key parameters. The locations were all roughly evenly spaced, located where the current could be physically felt, narrow enough for a tape measure to be reasonably stretched from bank to bank (less than 45 meters), and accessible via boat (with the government’s approval). In 2023, our research group gained permission to access further locations within the lake. Because of this, an additional four locations were chosen with the same parameters, leading to a total of 13 sites that stretched almost the entirety of the laguna.

The sites with their respective locations are shown in Table 1. Figure 2 depicts all the locations on a Google Earth map.

Table 1. Site Number, Site Name, Latitude, and Longitude of all Gauging Locations

<table>
<thead>
<tr>
<th>Site number</th>
<th>Name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Total width</th>
<th>Average depth</th>
<th>Maximum depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Elizabeth’s Channel</td>
<td>18.860545</td>
<td>88.165011</td>
<td>5 m</td>
<td>0.52 m</td>
<td>0.68 m</td>
</tr>
<tr>
<td>2</td>
<td>Chile Verde Flow South</td>
<td>18.858963</td>
<td>88.200402</td>
<td>5.5 m</td>
<td>0.75 m</td>
<td>0.95 m</td>
</tr>
<tr>
<td>3</td>
<td>Mangrove Bay</td>
<td>18.769208</td>
<td>88.295205</td>
<td>21 m</td>
<td>0.72 m</td>
<td>0.98 m</td>
</tr>
<tr>
<td>4</td>
<td>Eastern Channel South Mangrove Bay</td>
<td>18.721875</td>
<td>88.330494</td>
<td>31 m</td>
<td>1.36 m</td>
<td>1.92 m</td>
</tr>
<tr>
<td>5</td>
<td>Pirates’ Canal</td>
<td>18.671859</td>
<td>88.374855</td>
<td>30 m</td>
<td>0.49 m</td>
<td>0.88 m</td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Depth 1</td>
<td>Depth 2</td>
<td>Depth 3</td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
<td>------------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>6</td>
<td>Mariscal Northeast</td>
<td>18.668626</td>
<td>88.367517</td>
<td>34 m</td>
<td>1.28 m</td>
<td>1.55 m</td>
</tr>
<tr>
<td>7</td>
<td>East Laguna Mariscal</td>
<td>18.654930</td>
<td>88.368599</td>
<td>17 m</td>
<td>0.56 m</td>
<td>0.96 m</td>
</tr>
<tr>
<td>8</td>
<td>Mariscal Outlet</td>
<td>18.632708</td>
<td>88.384770</td>
<td>37 m</td>
<td>0.93 m</td>
<td>1.18 m</td>
</tr>
<tr>
<td>9</td>
<td>Hector’s Channel 1</td>
<td>18.631300</td>
<td>88.385368</td>
<td>16 m</td>
<td>0.52 m</td>
<td>0.65 m</td>
</tr>
<tr>
<td>10</td>
<td>Hector’s Channel 1B</td>
<td>18.616746</td>
<td>88.319916</td>
<td>9.5 m</td>
<td>0.53 m</td>
<td>0.73 m</td>
</tr>
<tr>
<td>11</td>
<td>Chaac River</td>
<td>18.623966</td>
<td>88.386588</td>
<td>42 m</td>
<td>0.80 m</td>
<td>1.07 m</td>
</tr>
<tr>
<td>12</td>
<td>Rapids</td>
<td>18.590169</td>
<td>88.436125</td>
<td>28 m</td>
<td>0.63 m</td>
<td>1.09 m</td>
</tr>
<tr>
<td>13</td>
<td>Xul-Ha</td>
<td>18.565264</td>
<td>88.447937</td>
<td>37 m</td>
<td>1.03 m</td>
<td>1.47 m</td>
</tr>
</tbody>
</table>
Figure 2. Locations of All Stream Gauging Sites
Stream gauging data was analyzed using midsection method guidelines (WMO, 2015). Calculation of streamflow using the midsection method is depicted in Equation 1 and Figure 3:

\[ q_n = v_n \left( \frac{b_n - b_{(n-1)}}{2} \right) d_n \]

where \( q_n \) is the partial discharge through a segment, \( v_n \) is the velocity of water through a segment, \( b_n \) is the distance from the initial segment, \( b_{(n-1)} \) is the distance from the initial segment to the preceding segment, and \( d_n \) is the depth of water at the segment measured.

Figure 3: Depiction of Midsection Method
Stream flows \((q_n)\) were measured using an OTT MF Pro (flow rate, \(v_n\)), vertical measuring rod (depth, \(d_n\)), and tape measure (width, \(b_n - b_{(n-1)}\)).

Gauging was performed in the following way: Location of interest was chosen. All personnel were transported to the location by boat. The boat and personnel set up downstream of the chosen location. A segment of the river or channel was chosen, and the width measured with a tape measure (Stanley 33-425 PowerLock™ 1” x 25’ Classic Tape Measure). Personnel were stationed every few meters to help keep the tape straight (Figure 4).

Figure 4. Personnel Holding the Tape Measure to Measure the Width of a Waterway

River channels were divided into 1- or 2-meter sections depending on the total width of the channel. Channels had a minimum of 10, with an average of 25, sections. A minimum of 10 sections is recommended for the total discharge or flow rate calculation by midsection method. An average of 25 sections was used, based on time constraints of this research project. The flowmeter (OTT HydroMet, OTT MF pro Meter, Portable with Case) with the attachable sensor (OTT HydroMet, OTT MF Pro Velocity and Depth Sensor, Cable 2 m) was placed at a height equal to one-third of the water depth from the bottom. Water depth \((d_n)\), flowmeter depth \((b_n - b_{(n-1)})\), and water
velocity \( (v_n) \) were measured at each section using the OTT MF Pro. It was key to make sure the OTT MF Pro completed a full cycle, approximately 30 seconds, before values were recorded. Discharge within each segment was calculated using Equation 1 and summed together to obtain total flow of the channel at each sampling location.

November 2022 and June and July 2023 were directly compared to observe the typical differences that can be expected from one year to the next. These dates were chosen because they contained the largest number of site locations. This helps establish differences that research can anticipate from one year to the next.

A water basin is a geological formation or structure that holds or collects water (Link et al., 2021). Laguna Bacalar was divided into three main basins, and Laguna Mariscal was considered a separate basin. The locations of the four basins are indicated in Figure 5. The residence time was calculated using Equation 2:

\[
\frac{(A \times D)}{(Q \times 3.15 \times 10^7 \text{ seconds per year})} \tag{2}
\]

where \( A \) is area in \( m^2 \), \( D \) is estimated depth based on bathymetry data, \( Q \) is inflow or outflow in \( m^3/s \), and \( 3.15 \times 10^7 \) is the conversion rate from seconds to years.

Bathymetry data were collected using the multibeam surveying method (Yunus et al., 2019). Bathymetry data were not available at certain locations because they had yet to be surveyed. On average, Laguna Bacalar is approximately 10 meters, so when no bathymetry data were available, 10 meters was used instead (Bnamericas, 2016).
Figure 5. Locations of Residence Time Water Basins
Results

Stream gauging of Laguna Bacalar first started in January 2015 at two locations (sites 11 and 12). Since then, stream gauging data have been collected at 11 additional locations: four added in March/April 2022 (sites 5, 7, 9, and 13), three in November 2022 (sites 3, 4, and 8), and four in June/July 2023 (sites 1, 2, 6, and 10). Table 2 depicts all the flow rate measurements (m$^3$/s) along with the month and year they were taken.

Table 2. Historical (January 2017–November 2022) and Current (June/July 2022) Stream Flow Data from Laguna Bacalar

<table>
<thead>
<tr>
<th>Site number</th>
<th>Jan 2017 (m$^3$/s)</th>
<th>Jan 2018 (m$^3$/s)</th>
<th>Jan 2019 (m$^3$/s)</th>
<th>May 2019 (m$^3$/s)</th>
<th>March and April 2022 (m$^3$/s)</th>
<th>Nov 2022 (m$^3$/s)</th>
<th>June and July 2023 (m$^3$/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.0</td>
<td>1.51</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.6</td>
<td>1.46</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>4.81</td>
<td>3.5</td>
<td>6.06</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.76</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>1.69</td>
<td>2.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.8</td>
<td>4.63</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>1.98</td>
<td>2.4</td>
<td>1.74</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>7.3</td>
<td>7.9</td>
<td>6.7</td>
<td>5.4</td>
<td>5.36</td>
<td>7.4</td>
<td>6.54</td>
</tr>
<tr>
<td>12</td>
<td>4.6</td>
<td>5.6</td>
<td>4.6</td>
<td>4.9</td>
<td>4.92</td>
<td>6.7</td>
<td>6.11</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.09</td>
<td>3.13</td>
</tr>
</tbody>
</table>

Maps of the flow pattern were created from the complete gauging efforts in 2022 and 2023. Figure 6 depicts the stream gauging data from November 2022, while Figure 7 depicts the data from June and July 2023. Table 3 directly compares the stream gauging data from November 2022 to that from June and July 2023. Tables 4 and 5 provide the residence time in the basins for November 2022 and for Summer 2023, respectively. Areas were obtained from Google Earth imagery. When possible, the inflow and outflow were recorded in m$^3$/s. Some locations do not include inflow or outflow because our team was unable to reach those locations via boat or did not have
permission from the government to access at the time. Residence time is calculated in days.

Figure 6. Stream Gauging Data from November 2022 (measurements in m$^3$/s)
Figure 7. Stream Gauging Data from June and July 2023 (measurements in m³/s)
Table 3. Stream Gauging Data Direct Comparison, November 2022 and Summer 2023

<table>
<thead>
<tr>
<th>Site number</th>
<th>November 2022 (m³/s)</th>
<th>June and July 2023 (m³/s)</th>
<th>Percent match</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2.0</td>
<td>1.51</td>
<td>75.5%</td>
</tr>
<tr>
<td>4</td>
<td>2.6</td>
<td>1.46</td>
<td>56.2%</td>
</tr>
<tr>
<td>5</td>
<td>3.5</td>
<td>6.06</td>
<td>173.1%</td>
</tr>
<tr>
<td>8</td>
<td>4.8</td>
<td>4.63</td>
<td>96.5%</td>
</tr>
<tr>
<td>9</td>
<td>2.4</td>
<td>1.74</td>
<td>72.5%</td>
</tr>
<tr>
<td>11</td>
<td>7.4</td>
<td>6.54</td>
<td>88.4%</td>
</tr>
<tr>
<td>12</td>
<td>6.7</td>
<td>6.11</td>
<td>91.2%</td>
</tr>
</tbody>
</table>

Table 4. Residence Time in Basins 1–3 and Mariscal Basin, November 2022

<table>
<thead>
<tr>
<th>Basin</th>
<th>Area (m²)</th>
<th>Depth (m)</th>
<th>Inflow (m³/s)</th>
<th>Outflow (m³/s)</th>
<th>Residence time based on inflow (days)</th>
<th>Residence time based on outflow (days)</th>
<th>Average residence time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.95 x 10⁶</td>
<td>23.8</td>
<td>5.35</td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>7.9 x 10⁶</td>
<td>10</td>
<td>6.2</td>
<td>3.5</td>
<td>147.5</td>
<td>261.2</td>
<td>204.4</td>
</tr>
<tr>
<td>3</td>
<td>4.6 x 10⁷</td>
<td>10</td>
<td></td>
<td>3.5</td>
<td>46.4</td>
<td></td>
<td>40.2</td>
</tr>
<tr>
<td>Mariscal</td>
<td>3.12 x 10⁶</td>
<td>4.5</td>
<td>3.5</td>
<td>4.8</td>
<td>33.9</td>
<td></td>
<td>40.2</td>
</tr>
</tbody>
</table>
Table 5. Residence time in Basins 1–3 and Mariscal Basin, Summer 2023

<table>
<thead>
<tr>
<th>Basin</th>
<th>Area (m²)</th>
<th>Depth (m)</th>
<th>Inflow (m³/s)</th>
<th>Outflow (m³/s)</th>
<th>Residence time based on inflow (days)</th>
<th>Residence time based on outflow (days)</th>
<th>Average residence time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.95 x 10^6</td>
<td>23.8</td>
<td>3.13</td>
<td></td>
<td>172</td>
<td>172</td>
<td>172</td>
</tr>
<tr>
<td>2</td>
<td>7.9 x 10^6</td>
<td>10</td>
<td>6.11</td>
<td>6.06</td>
<td>149.7</td>
<td>150.9</td>
<td>150.3</td>
</tr>
<tr>
<td>3</td>
<td>4.6 x 10^7</td>
<td>10</td>
<td>3.02</td>
<td></td>
<td>1762.9</td>
<td>1762.9</td>
<td>1762.9</td>
</tr>
<tr>
<td>Mariscal</td>
<td>3.12 x 10^6</td>
<td>4.5</td>
<td>6.06</td>
<td>6.39</td>
<td>26.8</td>
<td>25.4</td>
<td>26.1</td>
</tr>
</tbody>
</table>

Discussion

The objective of this study was to evaluate the flow rate and residence times of Laguna Bacalar. Residence time in water is an important concept in environmental science. It refers to the average amount of time a water molecule or substance, such as pollutants, spends in a particular system or body of water and can be an important indicator of overall health of a water system (Gilboa et al., 2022). The main comparisons were made between November of 2022 and Summer of 2023 because previous years lacked complete data sets.

It is important to account for seasonal changes, as seasonal shifts play an important role in the water supply and thus the flow rate and residence time of a given body of water. They can influence the movement, availability, and characteristics of water within a particular system. Previous research shows that residence time varies by 8.9 ± 0.4 days based on seasonality and precipitation (Van der Ent & Tuinenburg, 2017). If the residence times fall outside of this range between seasons, it can indicate other factors that may need to be considered.
Site 13 had nearly half the flow rate as site 12. This leaves a surprising 2.98 m³/s unaccounted for. This indicates that there must be another source of water coming into the lake and making its way to site 12. The unaccounted water is coming from groundwater input. A similar amount of unaccounted groundwater input was seen in March and April 2022 (Table 2). Future gauging of localized groundwater inputs will need to be performed before a definitive source can be determined.

Sites 8, 9, and 10 all flow into the Chaac River (site 11); their flow rates should therefore add up to that of site 11. This was clearly demonstrated in the stream gauging shown in Figure 6, where 4.63, 0.46, and 1.74 add up to 6.83. The 6.83 is nearly identical to the site 11 flow rate of 6.54, leading to a 95.8% match. This helps indicate that all the water flow is accounted for. The other stream gauging maps do not have site 10, so this observation was not possible.

Lake Mariscal is not at a steady state, with inflow and/or outflow rates varying outside of the anticipated seasonal differences. Pirates’ Canal (site 5) feeds directly into Mariscal Lake. If Laguna Mariscal is at steady state, Mariscal Outlet (site 8) would have the same flow rate. In November 2022, it was observed that Pirates’ Canal had an inflow of 3.5 and Mariscal’s outlet flow was 4.8. This would seem to indicate that some inflow is unaccounted for, but site 6 was not gauged that year. The missing inflow could be from that location. In summer 2023, Laguna Mariscal had an inflow of 6.06 and an outflow of 4.63 and 1.76, which adds to 6.39, indicating that all water was accounted for. Based on this, Mariscal’s residence time between these years differs. In November 2022, the average estimated residence time was 40.2 days, and in the summer of 2023, it was 26.1 days, leading to a difference of about 14 days due to seasonality, which is greater than 8.9 days. A possible explanation could be an additional water source (e.g., groundwater) making a new connection and flowing into Lake Mariscal. Additional stream gauging will need to be conducted in the following years.

In the summer of 2023, it was observed that water flowing into Lake Mariscal was coming from Pirates’ Canal at a flow rate of 6.06 and exiting through two different exit points. To the south, it exited with a flow rate of 4.63, and to the east, it exited with a flow rate of 1.76. In November 2022, the inflow from Pirates’ Canal was 3.5 and the outflow was 4.8, with no flow exiting to the east. This indicates that Mariscal Northeast becomes an active flow when Pirates’ Canal’s inflow is high. Although this seems to be the case, Mariscal Northeast has been measured only once, so further gauging is required for confirmation.

Despite all of this, this study has four independent measurements of residence time (the inflow and outflow for both November 2022 and Summer 2023). The overall average residence time for Lake Mariscal was 33.1 days, with a standard deviation of 8.3 days. Lake Mariscal’s inflow has been measured three separate times, with an average of 4.79 m³/s and a standard deviation of 1.0 m³/s. Mariscal’s outlet has been
measured seven times, with an average of 6.7 m³/s and a standard deviation of 0.9 m³/s.

Residence time was calculated for the four main water basins (locations shown in Figure 5). Residence times for basins 1 and 2 were extremely similar, with an approximate residence time of 0.5 year. This would mean that both basins are rapidly flushed, which is supported by the observation that the lake is extremely clear in those areas.

Basin 3, which is far larger than the other three basins, has a significantly larger residence time of 9.7 years, assuming the only outflow is to Chili Verde (1.51 m³/s). Another possibility is that the eastern channel also drains basin 3, adding an additional 1.51 m³/s from site 3. This cuts the residence time almost in half, with an alternative residence time of 4.8 years. Once again, this makes sense, given the small flow rates exiting the basin, although further study is required. Finally, Lake Mariscal had the shortest residence time at approximately 0.1 years. This makes sense given the small volume of the basin.

The lake’s residence time indicates how quickly water circulates through a water system. This research can help the government understand the long-term effects of tourism and, by extension, pollution on the lake. Based on this research, certain areas of the laguna will recover more quickly than other areas. Lake Mariscal has the shortest residence time, indicating that it can circulate and clear pollutants more quickly than other areas of the lake, such as basin 3, which has a significantly larger residence time.

Conclusion

This research project aimed to understand the flow rates and residence times within Laguna Bacalar. Within the lake, 13 areas of interest were identified and gauged using the midsection method. The lake was divided into four water basins, and the residence time was calculated for each using its respective inflows and outflows. Maps depicting the flow rates were constructed for both November 2022 and Summer 2023. It was found that most of the water flows into Laguna Bacalar from the south (sites 13 and 12) and exits through Pirates’ Canal (site 5). A small amount of water enters the laguna from the north exits through Chili Verde. The eastern channel flows southward toward Mariscal Lake and the Chaac River. Mariscal Lake is in a nonsteady state, with water flowing at different rates during different seasons. Overall, this project established a clear understanding of Laguna Bacalar’s hydrological system. Based on this study’s findings, it is clear that we have enhanced our understanding of ecosystem health and provided invaluable insights for effective research strategies, fostering a harmonious balance that benefits both the environment and human services, thereby
paving the way for future research that will in turn pave the way for a more sustainable and resilient future for Laguna Bacalar.
References


MODELING EMPLOYEE BURNOUT AND STRESS LEVELS USING GENERALIZED LINEAR MODELS

MADELINE S. NEELY, BUTLER UNIVERSITY
MENTOR: RASITHA R. JAYASEKARE

Abstract

In recent years, mental health has become a more prominent problem worldwide. One specific area of rising concern is the increase in the amount of stress and burnout that many people are experiencing. This research seeks to investigate the factors that are affecting stress and burnout as they pertain to the workplace environment. A series of multinomial logistic regression models and Poisson regression models are used to identify the factors affecting employees’ stress and burnout with respect to the workplace environment using the most recent Work, Family, and Health Study (WFHS) dataset from the Work, Family, and Health Network. This research found that important factors affecting stress and burnout include decision authority, job demands, and job satisfaction.

Introduction

Mental health has been a rising concern worldwide, as evidenced by the 13% increase in the incidence of mental health conditions within the past decade, according to the World Health Organization (n.d.). Additionally, the National Institute of Mental Health estimates that approximately one in five U.S. adults was living with some kind of mental illness in 2021 (n.d.). Stress and burnout at work have also been on the rise in recent years, and according to the American Institute of Stress (2019), 83% of employees in the United States suffer from work-related stress. These alarming statistics speak to the need for a better understanding of the causes of employee stress and burnout in order to improve mental health outcomes.

One area that has been heavily researched is the relationship between work and employee stress. This area of research is important for both employees’ personal health and well-being and for their employers’ overall success, as poor mental health can result in absenteeism, decreased work performance, bad attitudes and behaviors, and poor work relationships (Harnois et al., 2000). Because of these negative outcomes, many have begun to research what factors influence employee burnout and stress.
Christina Maslach and colleagues (2001) explain that some possible causes of burnout might include high job demands, a lack of autonomy, the absence of job resources, and individual factors such as personality characteristics and job attitudes. A case study in the UK came to similar conclusions, citing that a loss of control and feelings of powerlessness also contribute to employee stress (Harnois et al., 2000). Another source found sufficient findings in the literature that show the great impact that job satisfaction can have on workplace burnout (Friganović et al., 2019). Finally, another paper looked specifically at the difference between job stress and job burnout and discovered that they seem to be highly correlated with slightly different factors. Factors related to job strain were more highly correlated to stress, whereas burnout was more correlated with variables such as job dissatisfaction, a desire to quit, and emotions regarding perceived performance (Pines and Kinan, 2005).

This paper will focus on one particular study, the Work, Family, and Health Study (WFHS), which was led by the Work, Family, and Health Network (2018). This study was created to better understand how different workplace practices and policies influence employees’ work, family lives, and overall health and well-being. Many researchers have already used this dataset to gain a better understanding of the employee experience and how different factors can affect employees’ overall health and well-being. For example, using general linear models on the WFHS dataset, Lawson, Lee, and Maric (2021) focused on work-to-family conflict and how it could influence physical and mental health, discovering that increased conflict seems to be correlated with individuals who report poorer sleep and higher psychological distress. Fan et al. (2019) used group-based multitrajectory modeling on the WFHS dataset and found that workers in higher-strain environments have much worse long-term well-being outcomes than those in lower-strain environments. Vigoureux et al. (2020) have included stress as a predictor variable rather than a response variable in multilevel models to examine the relationship between sleep quantity and perceptions of stress. The researchers discovered that higher stress levels are strongly related to poorer metabolic health outcomes, showing that stress levels negatively affect employees’ mental and physical health.

Overall, much research has been done investigating employee burnout and stress. Many publications have been written on the analysis of the WFHS, but few have focused solely on understanding the factors that influence employee stress and burnout. This paper seeks to fill this gap and investigate the factors that affect employee stress and burnout by using the previously mentioned publicly accessible WFHS dataset.

About the Dataset

The WFHS data were gathered through randomized field experiments that assessed individuals in the information technology department of a Fortune 500
company (Work, Family and Health Network, 2018). Each participant in the study was asked a series of questions on topics such as the number of hours they work, how flexible their job is, what their work-life balance is like, how satisfying their job is, and how stressed and burned out their job makes them feel (Work, Family and Health Network, 2018). Two important variables to note in the WFHS dataset that will be explored throughout the remainder of this paper are the variables Psychological Distress and Emotional Burnout.

Psychological Distress

The first variable that this research focuses on is Psychological Distress. The measure used by the WFHS to quantify Psychological Distress is known as K6 and will be discussed in greater detail later in this paper. The distribution for this variable is shown in Figure 1. As can be seen in Figure 1, most respondents’ stress levels fell between 7 and 11, indicating that they fell within the range of low mental stress; however, it is also apparent that a fair number of respondents had scores ranging from 12 to 19, falling into the category of moderate mental stress, and a few fell into the category of severe psychological distress.

Employee Burnout

The second variable that this research will focus on is Burnout. This is a categorical variable that corresponds to the survey question “You feel burned out by your work. How often do you feel this way?” (Work, Family and Health Network, 2018). The bar chart in Figure 2 shows the frequency of different responses to this
question. In this chart, it appears that most respondents fell into the category of feeling burnout a few times a year to a few times a month. Quite a few people also answered that they felt burned out a few times a week; however, it’s important to note that there were fewer responses for the extremes of never or always feeling burned out.

Figure 2. Frequency of Respondent Burnout

Kessler 6 Psychological Distress Metric

The Kessler 6, or K6, is the measure used to evaluate the level of psychological distress of each employee. According to the WFHS (Work, Family and Health Network, 2018), “K6 is the most widely used mental health screening scale in the United States and has been used in numerous psychiatric and social epidemiology studies, including the National Household Survey on Drug Abuse.” The K6 metric is constructed by taking the sum of a respondent’s answers to six questions. These questions revolve around how often the respondent feels “nervous,” “hopeless,” “restless or fidgety,” “so depressed that nothing could cheer you up,” “that everything was an effort,” and “worthless” (Kim et al., 2016). The result is a discrete score ranging from 6 to 30. A score of 6 indicates that a respondent reported never feeling any of these feelings, and a 30 indicates that the respondent always felt each of these feelings.

It is easiest to understand the K6 scores in terms of low, medium, and high stress. According to Prochaska (2012), a score between 5 and 13 on a K6 scale ranging from 0 to 24 indicates moderate mental stress, and a score greater than 13 indicates severe psychological distress. When adjusting this to the K6 scale used in this study (from 6 to 30), the cutoff points become 11 (5 + 6) for moderate mental stress and 19
(13 + 6) for severe psychological distress. Table 1 shows these cutoff points for psychological distress.

Table 1. K6 Scores and Interpretations

<table>
<thead>
<tr>
<th>Score</th>
<th>Psychological Distress Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>No Psychological Distress</td>
</tr>
<tr>
<td>7-11</td>
<td>Low Psychological Distress</td>
</tr>
<tr>
<td>12-19</td>
<td>Moderate Psychological Distress</td>
</tr>
<tr>
<td>20-30</td>
<td>High Psychological Distress</td>
</tr>
</tbody>
</table>

Objective

The main objective of this research is to use the WFHS study data to identify how workplace practices and policies influence employee stress and burnout. The main objective will be achieved in two parts, by identifying the factors that affect employee stress and identifying the factors that affect employee burnout.

Methodology

Because of the types of response variables, we need two types of models to determine which factors are affecting employee burnout and employee stress. The response variable of the first model, Employee Stress, takes a discrete value between 6 and 30. Poisson random variables are frequently used to model counts with a minimum value of 0 and a maximum value that is unbounded (in theory) (Roback and Legler, 2021). Because Poisson regression predicts responses for count data, it is a good choice for modeling the Employee Stress response variable.

The response variable of the second model is Employee Burnout. This is a categorical variable that takes the values Never, A few times a year, Once a month, A few times a month, Once a week, A few times a week, and Every day. Multinomial logistic regression is used to model variables with a finite number of categorical outcomes. It does this by modeling the log odds of the outcomes in relation to a linear combination of predictor variables (UCLA: Statistical Consulting Group, 2021a, 2021b). Because
multinomial logistic regression can be used to classify variables with a set number of categorical responses, this is therefore a good fit for the Emotional Burnout response variable, which takes only one of seven possible values.

**Poisson Regression**

Poisson regression is an instance of the generalized linear model that is used when the response variable \( Y \) is a random variable that takes nonnegative integer values (Penn State: Eberly College of Science, 2016). Let \( Y \) be a Poisson random variable with mean \( \lambda \). Connecting the mean of the Poisson response variable and the linear combination of \( x_1, \ldots, x_k \) predictor variables using logarithm link function, the Poisson regression model can be written as

\[
\ln(\lambda) = \beta_0 + x_1\beta_1 + x_2\beta_2 + \ldots + x_k\beta_k .
\]

Then, by solving for \( \lambda \), we have the equivalent expression

\[
\lambda = \exp(\beta_0 + x_1\beta_2 + x_2\beta_2 + \ldots + x_k\beta_k) .
\]

The model coefficients, \( \beta_1, \ldots, \beta_k \), can be interpreted in the following way: A one-unit increase in the predictor variable \( x_i \) changes the log of the response variable by \( \beta_i \) units when all other predictor variables remain fixed.

**Multinomial Logistic Regression**

Because the response variable Emotional Burnout is a multiclass response variable, we use a multinomial logistic regression model. The multinomial logistic regression handles multiple classes, in comparison to the two classes handled similarly by the logistic regression model (Cheng et al., 2021). The multinomial logistic regression model reserves one response category as the base case and maps the log odds of other response categories compared to the base case category. Let the base case response category be denoted as \( Y = y_0 \) and the remaining categories be denoted as \( Y = y_1, \ldots, y_m \). Then the multinomial regression model looks like

\[
\log(P(Y = y_i) / P(Y = y_0)) = \beta_{0i} + x_1\beta_{1i} + x_2\beta_{2i} + \ldots + x_k\beta_{ki} ,
\]

where \( y_i = y_1, y_2, \ldots, y_m \).

The model coefficients can be interpreted in the following way: When \( x_i \) increases by one unit, the log odds of response variable category \( y_i \) relative to the base case \( y_0 \) would be expected to change by \( \beta_i \) units while keeping all other predictor variables fixed.
Data Preprocessing

Prior to modeling, some data preprocessing had to be performed to clean the data and make them usable for the chosen models. The first thing as a part of the preprocessing was to relabel to make the data easier to interpret quickly. For example, the variables were given new meaningful names that corresponded to the survey questions. In addition to this relabeling, the responses were recoded from their numerical form back into their original survey response answers. For example, if a variable had the value 1, we converted this back to strongly agree or every day, depending on the written response corresponding to that numeric value for the given question. After the data were relabeled, observations with missing values had to be removed because the machine learning algorithms that we selected do not perform well on data with missing values. Prior to the removal of observations with missing values, there were 3,684 observations, and after their removal, there were 2,794 observations. Another component of data preprocessing that we performed prior to modeling was outlier analysis. Other than missing values, however, no outliers were removed, because all data contained valid responses to the survey questions and we did not want to risk removing data that might give insights into important trends.

We then performed a multicollinearity analysis between the predictor variables to evaluate a standard regression assumption of no multicollinearity between independent variables. When a pair of predictor variables had a correlation coefficient greater than 0.5, one was removed after they were carefully analyzed. Then the principal component analysis (PCA) allowed us to reduce the number of predictor variables included in the regression models. The original goal was to include only the variables that directly corresponded to the survey response data, rather than the constructed variables that were the mean of these responses for each category. When the original variables were used, however, there were complications with building the models because many of the variables took similar values, resulting in infeasible model solutions. We therefore decided to focus only on the constructed quantitative variables given in the dataset.

The original dataset contained 21 quantitative variables. After a few of these attributes were removed because of multicollinearity, each model contained 18 attributes, making PCA no longer necessary. Table 2 shows the selected quantitative variables at the end of data preprocessing.
Table 2. Attributes Selected for Modeling

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Employee Stress</th>
<th>Employee Burnout</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActualWeeklyHours</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DecisionAuthority</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>EmotionalBurnout</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ExpectedWeeklyHours</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FamilyToWorkConflict</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>JobDemands</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>JobSatisfaction</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>LikelyToQuit</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>LowValueWork</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>NumberHoursInBed</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>OrganizationalCitizenship</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>PerceivedStress</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>PositiveWorkToFamily</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>RatePerformanceOthers</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>RatePerformanceSelf</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>SupervisorFamilySupport</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TimeAdequacy</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>WorkFamilyIssues</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>WorkHours</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Results

Identifying Factors that Affect Employee Stress

The first objective of this paper is to determine the factors affecting employee stress. The response variable, in this case, is Psychological Distress. This discrete variable identifies the Kessler 6 psychological distress score for each individual. This variable ranges from 6 to 29. Because this variable is a discrete count, a Poisson regression model works well. After removing the variables that cause multicollinearity within the dataset, the model is as shown below:

\[
\log(\text{AvgPsychologicalDistress}) = \beta_0 + \beta_1 \text{FamilyToWorkConflict} + \\
\beta_2 \text{NumberHoursInBed} + \beta_3 \text{PositiveWorkToFamily} + \beta_4 \text{RatePerformanceOthers} + \\
\beta_5 \text{PerceivedStress} + \beta_6 \text{WorkHours} + \beta_7 \text{RatePerformanceSelf} + \\
\beta_8 \text{VariableX} + \beta_9 \text{VariableY} + \epsilon
\]
After running this model, we evaluated the 95% confidence intervals for the coefficient estimate for each variable in the model. The 95% confidence intervals of the parameter estimate of the variables PerceivedStress, OrganizationalCitizenship, DecisionAuthority, and EmotionalBurnout did not contain zero in them, thus making them significant factors in the model. Table 3 shows all model coefficient estimates and confidence intervals.

Table 3. Psychological Distress Model Outputs

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Estimate</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept) *</td>
<td>1.653</td>
<td>1.3475</td>
<td>1.9581</td>
</tr>
<tr>
<td>ActualWeeklyHours</td>
<td>0.0008</td>
<td>-0.0006</td>
<td>0.0023</td>
</tr>
<tr>
<td>DecisionAuthority *</td>
<td>-0.0244</td>
<td>-0.047</td>
<td>-0.0017</td>
</tr>
<tr>
<td>EmotionalBurnout *</td>
<td>0.0343</td>
<td>0.0227</td>
<td>0.0459</td>
</tr>
<tr>
<td>ExpectedWeeklyHours</td>
<td>-0.0005</td>
<td>-0.0036</td>
<td>0.0026</td>
</tr>
<tr>
<td>FamilyToWorkConflict</td>
<td>0.0139</td>
<td>-0.0076</td>
<td>0.0354</td>
</tr>
<tr>
<td>JobDemands</td>
<td>-0.0207</td>
<td>-0.0428</td>
<td>0.0015</td>
</tr>
<tr>
<td>LikelyToQuit</td>
<td>0.0072</td>
<td>-0.0067</td>
<td>0.0211</td>
</tr>
<tr>
<td>LowValueWork</td>
<td>-0.0143</td>
<td>-0.0298</td>
<td>0.0013</td>
</tr>
<tr>
<td>NumberHoursInBed</td>
<td>0.0037</td>
<td>-0.0099</td>
<td>0.0174</td>
</tr>
<tr>
<td>OrganizationalCitizenship *</td>
<td>0.0404</td>
<td>0.0164</td>
<td>0.0644</td>
</tr>
<tr>
<td>PerceivedStress *</td>
<td>0.066</td>
<td>0.0604</td>
<td>0.0717</td>
</tr>
<tr>
<td>PositiveWorkToFamily</td>
<td>0.0125</td>
<td>-0.0102</td>
<td>0.0353</td>
</tr>
<tr>
<td>RatePerformanceOthers</td>
<td>0.0009</td>
<td>-0.0123</td>
<td>0.0142</td>
</tr>
<tr>
<td>RatePerformanceSelf</td>
<td>-0.0034</td>
<td>-0.0172</td>
<td>0.0104</td>
</tr>
<tr>
<td>SupervisorFamilySupport</td>
<td>0.0035</td>
<td>-0.0147</td>
<td>0.0217</td>
</tr>
<tr>
<td>TimeAdequacy</td>
<td>-0.0154</td>
<td>-0.041</td>
<td>0.0103</td>
</tr>
<tr>
<td>WorkFamilyIssues</td>
<td>-0.0085</td>
<td>-0.0264</td>
<td>0.0095</td>
</tr>
<tr>
<td>WorkHours</td>
<td>-0.0037</td>
<td>-0.0279</td>
<td>0.0204</td>
</tr>
</tbody>
</table>
Based on the parameter estimates, for PerceivedStress, a one-unit increase in PerceivedStress increases the log of PsychologicalDistress by 0.66 when all other predictor variables remain constant. For OrganizationalCitizenship, a one-unit increase in OrganizationalCitizenship increases the log of PsychologicalDistress by 0.0404 when all other predictor variables remain constant. For DecisionAuthority, a one-unit increase in DecisionAuthority decreases the log of PsychologicalDistress by 0.0244 when all other predictor variables remain constant. Finally, for EmotionalBurnout, a one-unit increase in EmotionalBurnout increases the log of PsychologicalDistress by 0.0343 when all other predictor variables remain constant.

Each of the significant variables mentioned above will be explored in greater detail in their own subsections. Each subsection references survey questions from the WFHS (Work, Family and Health Network, 2018).

Perceived Stress

The variable PerceivedStress refers to the sum of a respondent’s answers to four questions: (1) “During the past 30 days, how often have you felt that you were unable to control the important things in your life?” (2) “During the past 30 days, how often have you felt confident about your ability to handle your personal problems?” (3) “During the past 30 days, how often have you felt that things were going your way?” (4) “During the past 30 days, how often have you felt difficulties were piling up so high that you could not overcome them?”

As shown in the scatterplot in Figure 3, there is a moderately strong linear relationship ($r = 0.7$) between Perceived Stress and Psychological Distress. This same trend is also shown by the density graph, which breaks psychological distress into the four levels as described above. It is apparent that those with no psychological distress also tend to have lower perceived stress, and those with high psychological distress tend to have high perceived stress.
**Figure 3.** Perceived Stress and Psychological Distress Scatterplot (left) and Density Plot (right)

![Graphs showing perceived stress and psychological distress](image)

**Decision Authority**

The predictor variable Decision Authority is created by taking the mean of individual responses to three statements: (1) “Your job allows you to make a lot of decisions on your own.” (2) “On your job, you have very little freedom to decide how you do your work.” (3) “You have a lot of say about what happens on your job.”

In the scatter plot in Figure 4, there appears to be a weak negative relationship between psychological distress and decision authority. This means that as decision authority increases (i.e., when an individual feels they have more freedom to control their work), they have lower psychological distress. This same trend can be shown in the density plot, where it appears that those who have high psychological stress have lower median decision authority, and those with lower psychological stress have higher decision authority.
Emotional Burnout

Emotional Burnout is a different variable than the categorical response variable BurnedOut, which is discussed in the next section. Emotional Burnout is the mean of the respondents’ answers to the following questions, which do include the other response variable BurnedOut. The respondents were asked to answer three questions: (1) “You feel emotionally drained from your work. How often do you feel this way?” (2) “You feel burned out by your work. How often do you feel this way?” (3) “You feel used up at the end of the workday. How often do you feel this way?”

In the scatterplot in Figure 5, there appears to be a slightly positive relationship between employee burnout and psychological distress. This indicates that as feelings of emotional burnout increase, so do feelings of psychological distress. This is also much more clearly seen on the density plot, where those with high psychological stress have a much higher burnout compared to those with no or low psychological distress.
**Additional Results**

In addition to the variables described above, the variable OrganizationalCitizenship was also found to be statistically significant in the model for Psychological Distress. There appeared to be little to no correlation between Psychological Distress and OrganizationalCitizenship, however. This indicates that this might be an area for further research to understand better how organizational citizenship might affect psychological distress.

**Identifying Factors That Affect Employee Burnout**

The second objective of this paper is to determine the factors affecting employee burnout. The response variable in this case is BurnedOut. This categorical variable corresponds to the survey question “You feel burned out. How often do you feel this way?” (Work, Family and Health Network, 2018). The possible responses to this question include Never, A few times a year, Once a month, A few times a month, Once a week, A few times a week, and Every day. For this reason, multinomial logistic regression is a good model for the data.

Using the base case as $y_0 = Never$, the multinomial logistic regression model is as shown below:

$$
\log \left( \frac{P(Y = y_i)}{P(Y = y_0)} \right) = \beta_{0i} + x_1\beta_{1i} + x_2\beta_{2i} + \ldots + x_k\beta_{ki},
$$
where \( y_i = y_1, y_2, \ldots, y_m \),

where \( y_i = A \text{ few times a year, Once a month, A few times a month, Once a week, A few times a week, Every day} \)

Here, \( \beta_{ki} \) are the coefficients for the regression model. These coefficients correspond to each of the variables in the model, which are also listed below:

- SupervisorFamilySupport
- OrganizationalCitizenship
- RatePerformanceOthers
- RatePerformanceSelf
- TimeAdequacy
- DecisionAuthority
- FamilyToWorkConflict
- LikelyToQuit
- PerceivedStress
- ActualWeeklyHours
- JobDemands
- JobSatisfaction
- PositiveWorkToFamily
- NumberHoursInBed
- ExpectedWeeklyHours
- LowValueWork
- WorkHours
- WorkFamilyIssues

After running this model, we evaluated the 95% confidence intervals for each coefficient estimate in the model. This output can be seen in Figures 6–11.

Figure 6. Log Odds of \( A \text{ few times a year} \) Relative to \( \text{Never} \)
Figure 7. Log Odds of *Once a month* Relative to *Never*

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActualWeeklyHours</td>
<td>-0.0008</td>
<td>-0.0254</td>
<td>0.0238</td>
</tr>
<tr>
<td>DecisionAuthority</td>
<td>-0.0008</td>
<td>-0.4216</td>
<td>0.4201</td>
</tr>
<tr>
<td>ExpectedWeeklyHours</td>
<td>-0.0149</td>
<td>-0.0794</td>
<td>0.0496</td>
</tr>
<tr>
<td>FamilyToWorkConflict*</td>
<td>0.6002</td>
<td>0.1843</td>
<td>1.016</td>
</tr>
<tr>
<td>JobDemands*</td>
<td>0.9033</td>
<td>0.5366</td>
<td>1.2699</td>
</tr>
<tr>
<td>JobSatisfaction*</td>
<td>-0.6557</td>
<td>-1.1148</td>
<td>-0.1966</td>
</tr>
<tr>
<td>LikelyToQuit</td>
<td>-0.004</td>
<td>-0.2711</td>
<td>0.2632</td>
</tr>
<tr>
<td>LowValueWork*</td>
<td>0.2685</td>
<td>0.0642</td>
<td>0.5327</td>
</tr>
<tr>
<td>NumberHoursInBed</td>
<td>0.1631</td>
<td>-0.0557</td>
<td>0.3819</td>
</tr>
<tr>
<td>OrganizationalCitizenship*</td>
<td>-0.3695</td>
<td>0.101</td>
<td>0.7121</td>
</tr>
<tr>
<td>PerceivedStress*</td>
<td>0.2373</td>
<td>0.1323</td>
<td>0.3424</td>
</tr>
<tr>
<td>PositiveWorkToFamily*</td>
<td>0.4065</td>
<td>0.101</td>
<td>0.7121</td>
</tr>
<tr>
<td>RatePerformanceOthers</td>
<td>-0.0021</td>
<td>-0.2251</td>
<td>0.2209</td>
</tr>
<tr>
<td>RatePerformanceSelf</td>
<td>-0.2161</td>
<td>-0.4705</td>
<td>0.0383</td>
</tr>
<tr>
<td>SupervisorFamilySupport</td>
<td>-0.2478</td>
<td>-0.5901</td>
<td>0.0945</td>
</tr>
<tr>
<td>TimeAdequacy</td>
<td>0.1618</td>
<td>-0.2336</td>
<td>0.5572</td>
</tr>
<tr>
<td>WorkFamilyIssues</td>
<td>-0.1621</td>
<td>-0.4444</td>
<td>0.1202</td>
</tr>
<tr>
<td>WorkHours</td>
<td>0.3412</td>
<td>-0.0648</td>
<td>0.7471</td>
</tr>
</tbody>
</table>

Figure 8. Log Odds of *A few times a month* Relative to *Never*

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActualWeeklyHours</td>
<td>0.0172</td>
<td>-0.0082</td>
<td>0.0426</td>
</tr>
<tr>
<td>DecisionAuthority</td>
<td>-0.0241</td>
<td>-0.4516</td>
<td>0.4034</td>
</tr>
<tr>
<td>ExpectedWeeklyHours</td>
<td>-0.0112</td>
<td>-0.0749</td>
<td>0.0525</td>
</tr>
<tr>
<td>FamilyToWorkConflict</td>
<td>0.3903</td>
<td>-0.0307</td>
<td>0.8114</td>
</tr>
<tr>
<td>JobDemands*</td>
<td>1.5494</td>
<td>1.1701</td>
<td>1.9286</td>
</tr>
<tr>
<td>JobSatisfaction*</td>
<td>-0.9778</td>
<td>-1.443</td>
<td>-0.5127</td>
</tr>
<tr>
<td>LikelyToQuit</td>
<td>-0.0696</td>
<td>-0.3404</td>
<td>0.2012</td>
</tr>
<tr>
<td>LowValueWork*</td>
<td>0.4362</td>
<td>0.168</td>
<td>0.7043</td>
</tr>
<tr>
<td>NumberHoursInBed</td>
<td>0.0501</td>
<td>-0.171</td>
<td>0.2712</td>
</tr>
<tr>
<td>OrganizationalCitizenship*</td>
<td>-0.4552</td>
<td>-0.8315</td>
<td>-0.0789</td>
</tr>
<tr>
<td>PerceivedStress*</td>
<td>0.3393</td>
<td>0.2329</td>
<td>0.4458</td>
</tr>
<tr>
<td>PositiveWorkToFamily*</td>
<td>0.6323</td>
<td>0.3691</td>
<td>0.9555</td>
</tr>
<tr>
<td>RatePerformanceOthers</td>
<td>-0.0501</td>
<td>-0.2767</td>
<td>0.1765</td>
</tr>
<tr>
<td>RatePerformanceSelf</td>
<td>-0.3505</td>
<td>-0.6077</td>
<td>-0.0934</td>
</tr>
<tr>
<td>SupervisorFamilySupport</td>
<td>-0.296</td>
<td>-0.6413</td>
<td>0.0493</td>
</tr>
<tr>
<td>TimeAdequacy</td>
<td>0.0065</td>
<td>-0.3966</td>
<td>0.4097</td>
</tr>
<tr>
<td>WorkFamilyIssues</td>
<td>-0.379</td>
<td>-0.6660</td>
<td>-0.092</td>
</tr>
<tr>
<td>WorkHours</td>
<td>0.2373</td>
<td>-0.1761</td>
<td>0.6506</td>
</tr>
</tbody>
</table>
**Figure 9. Log Odds of *Once a week* Relative to *Never***

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActualWeeklyHours</td>
<td>0.0251</td>
<td>-0.0035</td>
<td>0.0538</td>
</tr>
<tr>
<td>DecisionAuthority</td>
<td>-0.0563</td>
<td>-0.5245</td>
<td>0.4119</td>
</tr>
<tr>
<td>ExpectedWeeklyHours</td>
<td>0.0105</td>
<td>-0.0569</td>
<td>0.078</td>
</tr>
<tr>
<td>FamilyToWorkConflict*</td>
<td>0.4731</td>
<td>0.0172</td>
<td>0.9289</td>
</tr>
<tr>
<td>JobDemands*</td>
<td>1.6823</td>
<td>1.2600</td>
<td>2.1047</td>
</tr>
<tr>
<td>JobSatisfaction*</td>
<td>-1.193</td>
<td>-1.6957</td>
<td>-0.6903</td>
</tr>
<tr>
<td>LikelyToQuit</td>
<td>-0.0221</td>
<td>-0.3188</td>
<td>0.2746</td>
</tr>
<tr>
<td>LowValueWork*</td>
<td>0.5417</td>
<td>0.2446</td>
<td>0.8388</td>
</tr>
<tr>
<td>NumberHoursInBed</td>
<td>0.1577</td>
<td>-0.0911</td>
<td>0.4064</td>
</tr>
<tr>
<td>OrganizationalCitizenship*</td>
<td>-0.4742</td>
<td>-0.9047</td>
<td>-0.0436</td>
</tr>
<tr>
<td>PerceivedStress*</td>
<td>0.3303</td>
<td>0.2139</td>
<td>0.4466</td>
</tr>
<tr>
<td>PositiveWorkToFamily*</td>
<td>0.7900</td>
<td>0.4061</td>
<td>1.1739</td>
</tr>
<tr>
<td>RatePerformanceOthers</td>
<td>0.0295</td>
<td>-0.2244</td>
<td>0.2834</td>
</tr>
<tr>
<td>RatePerformanceSelf*</td>
<td>-0.436</td>
<td>-0.7164</td>
<td>-0.1556</td>
</tr>
<tr>
<td>SupervisorFamilySupport</td>
<td>-0.3272</td>
<td>-0.7028</td>
<td>0.0485</td>
</tr>
<tr>
<td>Time Adequacy</td>
<td>-0.2137</td>
<td>-0.6665</td>
<td>0.2391</td>
</tr>
<tr>
<td>WorkFamilyIssues*</td>
<td>-0.3718</td>
<td>-0.6955</td>
<td>-0.048</td>
</tr>
<tr>
<td>WorkHours</td>
<td>0.1011</td>
<td>-0.3582</td>
<td>0.5605</td>
</tr>
</tbody>
</table>

**Figure 10. Log Odds of *A few times a week* Relative to *Never***

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate</th>
<th>2.50%</th>
<th>97.50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActualWeeklyHours</td>
<td>0.0174</td>
<td>-0.0097</td>
<td>0.0446</td>
</tr>
<tr>
<td>DecisionAuthority</td>
<td>-0.231</td>
<td>-0.6825</td>
<td>0.2204</td>
</tr>
<tr>
<td>ExpectedWeeklyHours</td>
<td>0.0092</td>
<td>-0.0561</td>
<td>0.0746</td>
</tr>
<tr>
<td>FamilyToWorkConflict*</td>
<td>0.5358</td>
<td>0.0946</td>
<td>0.9769</td>
</tr>
<tr>
<td>JobDemands*</td>
<td>2.4948</td>
<td>2.0795</td>
<td>2.9101</td>
</tr>
<tr>
<td>JobSatisfaction*</td>
<td>-1.5565</td>
<td>-2.0454</td>
<td>-1.0676</td>
</tr>
<tr>
<td>LikelyToQuit</td>
<td>0.032</td>
<td>-0.2542</td>
<td>0.3181</td>
</tr>
<tr>
<td>LowValueWork*</td>
<td>0.3856</td>
<td>0.0986</td>
<td>0.6725</td>
</tr>
<tr>
<td>NumberHoursInBed</td>
<td>0.0176</td>
<td>-0.2197</td>
<td>0.2549</td>
</tr>
<tr>
<td>OrganizationalCitizenship*</td>
<td>-0.6496</td>
<td>-1.0643</td>
<td>-0.235</td>
</tr>
<tr>
<td>PerceivedStress*</td>
<td>0.386</td>
<td>0.2733</td>
<td>0.4987</td>
</tr>
<tr>
<td>PositiveWorkToFamily*</td>
<td>0.5713</td>
<td>0.2112</td>
<td>0.9315</td>
</tr>
<tr>
<td>RatePerformanceOthers</td>
<td>-0.0328</td>
<td>-0.2748</td>
<td>0.2092</td>
</tr>
<tr>
<td>RatePerformanceSelf*</td>
<td>-0.3749</td>
<td>-0.6467</td>
<td>-0.103</td>
</tr>
<tr>
<td>SupervisorFamilySupport*</td>
<td>-0.373</td>
<td>-0.7368</td>
<td>-0.0091</td>
</tr>
<tr>
<td>Time Adequacy</td>
<td>-0.2776</td>
<td>-0.7177</td>
<td>0.1625</td>
</tr>
<tr>
<td>WorkFamilyIssues*</td>
<td>-0.6492</td>
<td>-0.963</td>
<td>-0.3354</td>
</tr>
<tr>
<td>WorkHours</td>
<td>0.283</td>
<td>-0.1609</td>
<td>0.7268</td>
</tr>
</tbody>
</table>
Figure 11. Log Odds of *Every day* Relative to *Never*

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimate</th>
<th>2.50%</th>
<th>97.50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActualWeeklyHours</td>
<td>0.0225</td>
<td>-0.0118</td>
<td>0.0568</td>
</tr>
<tr>
<td>Decision Authority*</td>
<td>-0.5722</td>
<td>1.1011</td>
<td>-0.0433</td>
</tr>
<tr>
<td>ExpectedWeeklyHours</td>
<td>0.0264</td>
<td>-0.0474</td>
<td>0.1001</td>
</tr>
<tr>
<td>FamilyToWorkConflict</td>
<td>0.3709</td>
<td>-0.1568</td>
<td>0.8985</td>
</tr>
<tr>
<td>JobDemands*</td>
<td>3.0772</td>
<td>2.536</td>
<td>3.6183</td>
</tr>
<tr>
<td>JobSatisfaction*</td>
<td>-1.9653</td>
<td>-2.5347</td>
<td>-1.396</td>
</tr>
<tr>
<td>LikelyToQuit</td>
<td>0.004</td>
<td>-0.3458</td>
<td>0.3358</td>
</tr>
<tr>
<td>LowValueWork</td>
<td>0.3018</td>
<td>-0.0653</td>
<td>0.6689</td>
</tr>
<tr>
<td>NumberHoursInBed</td>
<td>0.0092</td>
<td>-0.297</td>
<td>0.3153</td>
</tr>
<tr>
<td>OrganizationalCitizenship*</td>
<td>-0.069</td>
<td>-0.6263</td>
<td>-0.2028</td>
</tr>
<tr>
<td>PerceivedStress*</td>
<td>0.4997</td>
<td>0.3622</td>
<td>0.6373</td>
</tr>
<tr>
<td>PositiveWorkToFamily*</td>
<td>0.6195</td>
<td>0.1181</td>
<td>1.1209</td>
</tr>
<tr>
<td>RatePerformanceOthers</td>
<td>-0.0389</td>
<td>-0.3378</td>
<td>0.2601</td>
</tr>
<tr>
<td>RatePerformanceSelf*</td>
<td>-0.502</td>
<td>-0.83</td>
<td>-0.1739</td>
</tr>
<tr>
<td>SupervisorFamilySupport*</td>
<td>-0.6351</td>
<td>-1.0674</td>
<td>-0.2028</td>
</tr>
<tr>
<td>TimeAdequacy</td>
<td>-0.4906</td>
<td>-1.0742</td>
<td>0.093</td>
</tr>
<tr>
<td>WorkFamilyIssues*</td>
<td>-0.6672</td>
<td>-1.088</td>
<td>-0.2464</td>
</tr>
<tr>
<td>WorkHours</td>
<td>0.1019</td>
<td>-0.4647</td>
<td>0.6685</td>
</tr>
</tbody>
</table>

Our criterion for a predictor variable to be significant for a burnout frequency level (e.g., log odds of *A few times a year* relative to *Never*) follows: We will call a predictor variable for a burnout level significant if zero is not in the 95% confidence interval estimate for the predictor’s corresponding $\beta$ parameter.

Using this criterion, the predictor variables that are significant for at least one of the burnout frequency levels are FamilyToWorkConflict, JobDemands, JobSatisfaction, LowValueWork, OrganizationalCitizenship, PerceivedStress, PositiveWorkToFamily, RatePerformanceSelf, SupervisorFamilySupport, and WorkFamilyIssues. Figure 12 shows a summary of the significant results.
Following are a few examples of interpretations of some of the variables that are significant in the model. For PositiveWorkToFamily, when PositiveWorkToFamily increases by 1 unit, the log odds for burned out a few times a year relative to never burned out would be expected to increase by 0.544 while keeping all other variables constant. For FamilyToWorkConflict, when FamilyToWorkConflict increases by 1 unit, the log odds for burned out once a month relative to never burned out would be expected to increase by 0.6002 while keeping all other factors fixed. For OrganizationalCitizenship, when OrganizationalCitizenship increases by 1 unit, the log odds for burned out every day relative to never burned out would be expected to decrease by 0.069 while keeping all other factors fixed. All other coefficients can be interpreted similarly.

Now we will explore in greater detail each of the variables that was found to be significant in the model. Investigating these variables further leads to a greater

<table>
<thead>
<tr>
<th>Log Odds Of...</th>
<th>Predictor Variable</th>
<th>Log Odds Of...</th>
<th>Predictor Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Few Times a Year Relative to Never</td>
<td>JobDemands PerceivedStress PositiveWorkToFamily RatePerformanceSelf</td>
<td>Once a Week Relative to Never Continued</td>
<td>OrganizationalCitizenship PerceivedStress PositiveWorkToFamily RatePerformanceSelf WorkFamilyIssues</td>
</tr>
<tr>
<td>Once a Month Relative to Never</td>
<td>FamilyToWorkConflict JobDemands JobSatisfaction LowValueWork OrganizationalCitizenship PerceivedStress PositiveWorkToFamily</td>
<td>A Few Times a Week Relative to Never</td>
<td>FamilyToWorkConflict JobDemands JobSatisfaction LowValueWork OrganizationalCitizenship PerceivedStress PositiveWorkToFamily RatePerformanceSelf SupervisorFamilySupport WorkFamilyIssues</td>
</tr>
<tr>
<td>A Few Times a Month Relative to Never</td>
<td>JobDemands JobSatisfaction LowValueWork OrganizationalCitizenship PerceivedStress PositiveWorkToFamily RatePerformanceSelf WorkFamilyIssues</td>
<td>Everyday Relative to Never</td>
<td>DecisionAuthority JobDemands JobSatisfaction OrganizationalCitizenship PerceivedStress</td>
</tr>
<tr>
<td>Once a Week Relative to Never</td>
<td>FamilyToWorkConflict JobDemands JobSatisfaction LowValueWork</td>
<td></td>
<td>PositiveWorkToFamily RatePerformanceSelf SupervisorFamilySupport WorkFamilyIssues</td>
</tr>
</tbody>
</table>
understanding of how different work conditions and environments are related to emotional burnout.

Perceived Stress

Perceived Stress is the same predictor variable as described in the Psychological Distress section. In the density plot in Figure 13, we see a clear relationship between the variables BurnedOut and Perceived Stress. Those who never or rarely feel burned out seem to have much lower perceived stress as compared to those who feel burned out every day. This indicates that there is likely some kind of relationship between burnout experience and the quantity of perceived stress that someone feels.

Figure 13. Density Plot of Perceived Stress and BurnedOut
Job Demands

Job Demands is a predictor variable that corresponds to the mean of the respondents’ responses to the following three statements: (1) “You do not have enough time to get your job done.” (2) “Your job requires very fast work.” (3) “Your job requires very hard work.”

The density plot in Figure 14 again shows a clear relationship between the predictor variable of job demands and the response variable of feeling burned out. Those with lower burnout tend also to have lower job demands, and those with higher burnout tend to have higher job demands. This makes sense and could mean that decreasing job demands by ensuring employees have enough time to get their work done could lead to decreased burnout experience.
Job Satisfaction

The variable JobSatisfaction corresponds to the mean of the respondents’ responses to the following statements: (1) “In general, you like working at your job.” (2) “In general, you are satisfied with your job.” (3) “You are generally satisfied with the kind of work you do in this job.”

There appear to be a few key trends in the density plot in Figure 15. Those who experience burnout every day tend to have a much lower median level of job satisfaction, although there is also a much larger spread of job satisfaction for those who experience burnout every day. Another interesting thing to note is that those who never or rarely experience burnout have a much higher level of job satisfaction. This means if a company could determine how to help its employees achieve greater job satisfaction, the company could also potentially decrease how frequently employees feel burned out.
Low-Value Work

The variable LowValueWork corresponds to the mean of the responses to the following statements: (1) “You work on unnecessary things.” (2) “You spend time in unproductive meetings.”

The density plot in Figure 16 shows a clear trend that those who rate their work as having lower value (meaning they feel their time is frequently wasted) often also experience burnout more frequently, whereas those who do not feel like time is wasted tend to have a lesser frequency of experiencing burnout. This could lead to some simple solutions for employers. For example, one idea might be limiting meetings to only necessary ones and decreasing busy work to help decrease employee burnout.

Figure 16. Density Plot for LowValueWork and BurnedOut
Work-Family Issues

The WorkFamilyIssues variable corresponds to the mean of the responses to the following three statements: (1) “In your workplace, employees are expected to take time away from their family or personal lives to get their work done.” (2) “In your workplace, employees are expected to put their families or personal lives second to their jobs.” (3) “In your workplace, employees are expected to make work their top priority.”

The density plot in Figure 17 shows a few clear trends. For example, it shows that those with higher burnout experience tend to have lower work-family issues. Because this was a reverse-coded variable, this means that they feel as though work takes priority over their family or personal lives very frequently, whereas those who rarely or never experience burnout tend to have higher work-family-issue scores, indicating that they don’t often feel that work must be their top priority over their family and personal lives.

Figure 17. Density Plot for WorkFamilyIssues and BurnedOut
Additional Findings

In addition to the results discussed above, a few other variables were found to be statistically significant within the employee burnout model at the 5% significance level. These variables were PositiveWorkToFamily, OrganizationalCitizenship, RatePerformanceSelf, FamilyToWorkConflict, and SupervisorFamilySupport. When visualizations were created and the relationship between these variables and emotional burnout were investigated, however, no obvious meaningful trends or patterns were found. This indicates that in the future it will be important to dive deeper into the correlation between these variables and burnout, as they could help provide more insight on what affects burnout.

Conclusion

This research sought to identify the factors affecting psychological distress and burnout using a Poisson regression model and a multinomial logistic regression model, respectively. In addition to the models discussed in this paper, various other models were also tested throughout this research. Because of the number of categorical variables in the dataset, we suspected that the model might involve random effects. To incorporate mixed effects in a multinomial setting, we first used the MCMCglmm package, which fits multivariate generalized linear mixed models using Markov chain Monte Carlo techniques (Hadfield, 2010). This was a significantly more complicated model, however, and did not find any random effects to be significant in the model, regardless of the variable selected as the random effects variable. For this reason, we used the simpler multinomial logistic regression model. We also explored mixed-effects Poisson regression for the psychological distress response variable; however, again, no random effects were found to be significant in the model. For this reason, the simpler Poisson regression model was selected.

The Poisson regression model for psychological distress concluded that PerceivedStress, OrganizationalCitizenship, DecisionAuthority, and EmotionalBurnout were significant based on coefficient estimates via 95% confidence intervals. The multinomial logistic regression model for emotional burnout concluded that PositiveWorktoFamily, OrganizationalCitizenship, PerceivedStress, JobDemands, RatePerformanceSelf, JobSatisfaction, LowValueWork, WorkFamilyIssues, FamilyToWorkConflict, and SupervisorFamilySupport were significant for at least one level of the multinomial values based on coefficient estimates via 95% confidence intervals.

An interesting finding is that the variables PerceivedStress and OrganizationalCitizenship were found to be significant in both the psychological distress model and emotional burnout model. This indicates that there is at least some overlap between the factors that affect stress and burnout. Additionally, emotional
burnout and perceived stress were observed to be significant variables in determining (respectively) psychological distress and burnout; thus, there is a clear relationship between psychological stress and emotional burnout.

This information can be used to help employers create work environments where employee burnout and stress are minimized. Based on the significant factors, some actionable items for employers would be to ensure that employees’ work is valued and that employers effectively use their employees' time, as this corresponds to the low-value-work factor. In addition, properly training managers to be supportive supervisors who understand and promote work-life balance pertains to the factors of work-family issues, family-to-work conflict, and supervisor family support. Finally, employers should ensure that they effectively balance the workloads assigned to employees, reducing unnecessary job demands and therefore hopefully reducing employee burnout and stress as well.

This research not only is beneficial for employers but can also be helpful for those looking to join the workforce or switch careers, as this gives insights on types of companies an individual might want to look for. First, they will want a job that brings them satisfaction, followed by a leader who will support their career growth and their work-life balance in a supportive community of workers.

Research on psychological distress and emotional burnout are useful topics and can further be extended into many other avenues. One such is cluster analysis, which uncovers natural groupings of the data, potentially revealing more relationships between different variables.

Finally, this research could also be expanded upon with examination of stress and burnout in different environments. For example, a similar study focused on college or high school students could help provide insights into how different learning environments and teachers or professors can affect student mental health. This research provides valuable insights into factors affecting employee stress and burnout and opens the door to future areas of investigation to help fight the ever-increasing mental health concerns in the world today.
References


ENHANCED BREAST CANCER TUMOR CLASSIFICATION USING MOBilenetV2: A DETAILED EXPLORATION ON IMAGE INTENSITY, ERROR MITIGATION, AND STREAMLIT-DRIVEN REAL-TIME DEPLOYMENT

AADITYA SURYA, UNIVERSITY OF SOUTHERN CALIFORNIA
ADITYA KESHARY SHAH, ROCHESTER INSTITUTE OF TECHNOLOGY
SUBASH TARUN SASIKUMAR, STATE UNIVERSITY OF NEW YORK PLATTSBURGH
JARNELL KABORE, NORTHEASTERN UNIVERSITY
MENTOR: SWAPNIL SNEHAM

Abstract

This research introduces a sophisticated transfer learning model based on Google’s MobileNetV2 for breast cancer tumor classification into normal, benign, and malignant categories, utilizing a dataset of 1576 ultrasound images (265 normal, 891 benign, 420 malignant). The model achieves an accuracy of 0.82, precision of 0.83, recall of 0.81, ROC-AUC of 0.94, PR-AUC of 0.88, and MCC of 0.74. It examines image intensity distributions and misclassification errors, offering improvements for future applications. Addressing dataset imbalances, the study ensures a generalizable model. This work, using a dataset from Baheya Hospital, Cairo, Egypt, compiled by Walid Al-Dhaybany and colleagues (2020), emphasizes MobileNetV2’s potential in medical imaging, aiming to improve diagnostic precision in oncology. Additionally, the paper explores Streamlit-based deployment for real-time tumor classification, demonstrating MobileNetV2’s applicability in medical imaging and setting a benchmark for future research in oncology diagnostics.

Keywords: MobileNetV2, image intensity error mitigation, Streamlit deployment, transfer learning, deep learning in oncology, ultrasound imaging, classification accuracy, convolutional neural networks (CNNs), medical image processing

Introduction

Breast cancer is a particular variant of carcinoma distinguished by its origin in the lining of breast tissue, leading to tumor formation with potential metastatic tendencies. It is typically shown to occur predominantly in women and holds the unfortunate title as the second most common adenocarcinoma within this demographic. The World Health Organization (WHO, 2021) suggests that more than
99% of breast cancer diagnoses concern women, highlighting the severe gender-centric nature of this disease.

Amidst this regrettable backdrop, it is noted that it is of vital necessity to timely diagnose such cases of cancer promptly, as the survival rates of patients are contingent largely on how expeditiously a diagnosis can be found; hence, the paramount purpose of this study is to create a model that robustly and efficiently produces correct classifications for a given ultrasound scan.

The integration of deep learning into breast cancer research has shown promise that AI models can offer enhanced patient examination procedures and classify diagnoses with both speed and accuracy (Al-Dhabyani et al., 2020; Stark et al., 2019). Our study leverages a dataset of breast ultrasound images meticulously curated by Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Aly Fahmy (2020). This invaluable resource proves to be the foundation upon which the following machine learning models will be trained, tested, and evaluated.

The particular dataset utilized for this research has classified types of breast cancer tumors into three categories: normal, benign, and malignant. A classification of normal indicates the absence of any tumor detected in the image, hence indicating healthy breast tissue. A classification of benign, in comparison, indicates that a growth is present but is noncancerous and does not possess the ability to spread or to invade surrounding tissue (Al-Dhabyani et al., 2020). The last classification presented is malignant, the most concerning of the three, as it indicates cancerous growths that have the potential to invade nearby tissue and metastasize to distant organs (Al-Dhabyani et al., 2020).
Table 1. Classification and Description of Tumor Types

<table>
<thead>
<tr>
<th>Class</th>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal</td>
<td>Corresponds to healthy tissue without any abnormalities or tumor growth</td>
</tr>
<tr>
<td>1</td>
<td>Benign</td>
<td>Refers to a noncancerous tumor that does not invade nearby tissues or spread to other parts of the body</td>
</tr>
<tr>
<td>2</td>
<td>Malignant</td>
<td>Represents cancerous tumors that can invade nearby tissues and spread to other parts of the body</td>
</tr>
</tbody>
</table>

The data are utilized through Google’s MobileNetV2 architecture for classifying breast cancer tumors into normal, benign, and malignant categories as shown in Table 1. An important acknowledgment to make are the limitations: hardware and software complexity. The hardware available is not at the forefront of technology and is rather commercial, ultimately restricting the model’s full capabilities. Additionally, the model complexity was affected through the same issue as the memory allocation problem cannot be overlooked. A lower degree of memory leads to smaller batch sizes and directly affects the model’s ability to recognize more abstract elements in the data. The last limitation was the low quantity of data available to train the model.

Intended Contributions and Research Question

The paper wants to answer the question, or rather the problem, of how the general public is supposed to access machine learning models. Up until the recent exponential increase in popularity of artificial intelligence and machine learning due to the efforts of OpenAI, many models were left in the hands of giant corporations who were the only ones looking into profiting off of them. Even now, it is still difficult to find a trusted model to use to check for health purposes. The paper is looking into this issue and trying to offer up an idea of what could be done with more time and proper funding.
The full extent of the paper reaches an online deployment of the model bridging the gap between the academic bubble and the general public. The implications of the model lie closer to the healthcare sector than to the machine learning sector, as it has a wider impact on cancer research and solutions. Because of this, the goal is to make a very user-friendly model for people to be able to use as a resource if need be. It is hoped that this will lower the number of false diagnoses across the nation and get people the help they need early on rather than later.

Literature Review

Foundational Techniques and Approaches in Breast Cancer Diagnosis

G. F. Stark and colleagues (2019) embarked on a comprehensive journey, analyzing an array of machine learning models to enhance the accuracy of breast cancer diagnosis. Their work positioned the traditional Gail model, a cornerstone in breast cancer risk assessment, against modern machine learning frameworks. This comparative evaluation is especially significant as it provides an empirical benchmark for model performance, emphasizing the evolution and transformative potential of AI in medical diagnostics. Researchers can draw from the methodology of Stark et al. to juxtapose traditional and modern diagnostic methods, ensuring that advancements in AI-driven systems are grounded in established medical knowledge.

Ahmad et al. (2013) focused on the intersection of data mining techniques and breast cancer prediction. Their research emphasized the profound capabilities of data mining, offering a systematic methodology to extract relevant patterns and insights from vast datasets. By merging these techniques with machine learning models, they paved the way for a more sophisticated understanding of breast cancer diagnostics. For researchers aiming to leverage breast cancer datasets, especially intricate ones like ultrasound datasets, integrating such data extraction techniques can lead to more enriched and informed model training processes.

The intersection of medical knowledge and machine learning models is the core of the paper, which answers the question of why papers about traditional medical techniques and modern machine learning techniques were cited. Transfer learning techniques were a great help in achieving this, as they allowed for the use of a pretrained model of a similar task (identifying tumors in breast cancer) as the starting point for a new model. This made the model already have an idea of what to do and how to do it.

Advanced Classification and Transfer Learning Techniques

Gulzar (2023) highlights an implementation of transfer learning using MobileNetV2 for fruit image classification. Though this may not be a direct
implementation of transfer learning within the context of the medical field, it promptly addresses and describes how convolutional neural networks can be enhanced by pretrained models to accurately classify images. The paper utilized a dataset of 26,149 images across 40 classes of fruits and introduced a modified MobileNetV2 incorporating a customized head to increase model robustness. The modified model leveraged transfer learning from MobileNetV2 and achieved an accuracy of 99%. When compared with other architectures, the transfer learning model consistently outperforms in image classification.

Another instance of transfer learning is seen in the work of Praba Hridayami and colleagues (2020), who effectively showcased the merits of utilizing the pretrained VGG16 model within their convolutional neural network (CNN) framework. Their approach, resulting in a high accuracy rate, underpins the importance of leveraging transfer learning techniques in medical imaging classification. Given the intricate patterns and nuances present in breast ultrasound imaging, integrating deep, pretrained models can offer a solid foundation, optimizing the model’s ability to distinguish among benign, malignant, and normal patterns. Hridayami and colleagues’ work underscores the need to balance depth and adaptability in machine learning architectures for high-stakes tasks such as cancer detection.

A deeper implementation of a cancer-prediction model can be seen in Mi and colleagues’ (2019) multiclass classification study of breast digital pathology, which also utilized deep learning for cancer prediction. This study in particular stands out, as it diverges from a simple binary classification model (such as normal vs. tumor or benign vs. malignant) and instead introduces a more profound multiclass classification system. Through employing a dual-stage architecture, the study classifies breast digital pathology images into four categories: normal tissue, benign lesion, ductal carcinoma in situ, and invasive carcinoma. This research proves to be a fundamental basis to building a multiclass classification system that is pertinent to the study at hand.

The machine learning model is thus able to utilize transfer learning techniques, but in order to make sure that the model does not go astray, evaluation metrics were relied upon. The evaluation metrics ensure the model’s accuracy and repeatedly check whether the model is accurately identifying benign, normal, or malignant tumors.

Evaluation Metrics and their Importance

Meysam Vakili and colleagues (2020) provide an analysis of different evaluation metrics for classification algorithms. Their paper serves as a meta-analysis, examining a plethora of metrics that encompass precision, recall, F1-score, accuracy, confusion matrices, and receiver operating characteristic area under the
curve (ROC-AUC) scores. In the notably instructive section 3.3 of their paper, the authors state the definitions and formulas for each metric, as well as their respective benefits and detriments in terms of evaluating classification models. For the purpose of this study, Vakili et al.’s exhaustive overview is invaluable as the choice of an evaluation metric, for the model presented below can heavily influence the interpretation of the model’s efficacy.

In addition to ROC-AUC scores, another metric of accuracy was researched. This creates a more holistic evaluation score, as multiple metrics will help discern unique limitations within the model. Chicco and Jurman (2020), in their paper comparing the Matthews correlation coefficient (MCC) to other relevant metrics, show the advantages that MCC has in evaluating classification models. Their paper compares MCC critically with F1 score and accuracy, which are often the most commonly used metrics for CNN models. Though accuracy and F1 score have been popular over the past few decades, the paper shows how often they present overoptimistic results, especially in the face of imbalanced datasets. The study describes how MCC, in contrast, provides a significantly more robust evaluation by accounting for true positives, false negatives, true negatives, and false positives while adjusting proportionally for positive and negative elements in the dataset. Given the study, leveraging MCC as an evaluation metric offers a rigorous and holistic understanding in classification studies, ensuring the authenticity of diagnostic outcomes.

Optimization and Hyperparameter Tuning in Neural Networks

A hyperparameter is an external configuration variable that is used to manage the training of machine learning models. Examples of this include the number of nodes/layers in a neural network and the number of branches in a decision tree. Hyperparameters determine the significant features, such as the model architecture, learning rate, and model complexity.

Because of the importance of hyperparameters, it is imperative to understand the current breakthroughs in optimizing and manipulating them for maximum effect. Saleh et al. (2022), in their research in different deep learning approaches toward breast cancer diagnosis, utilize Keras-Tuner in optimizing their architecture. The paper offers an advanced methodology using an optimized recurrent neural network (RNN). The Keras-Tuner in the paper is described as being chosen for its flexibility, user-friendliness, and streamlined manner in optimizing hyperparameters without the cumbersome trial-and-error process. In this study, the tuner optimized different dropout rates between 0.1 and 0.9 and showed a significant improvement in diagnostic accuracy with the optimizations.
Methodology

Data Collection

The data source is a series of photos and data collected in 2018 about baseline breast ultrasounds taken from women between the ages of 25 and 75. The dataset comprises 1576 images in PNG format, classified into three distinct categories: normal, malignant, and benign. A normal scan indicates the absence of any abnormal growths in the breast. A scan classified as malignant indicates the presence of harmful growths, and a scan classified as benign signifies the detection of harmless growths.

To aid in understanding the distribution of the dataset, Figure 1 presents a pie chart detailing the numerical and percentage distribution of each class, specifically, the 265 normal instances, 891 benign instances, and 421 malignant instances. This visualization is pivotal for our paper, as it underlines the inherent class distribution and aids in understanding potential challenges related to class imbalances. Addressing and understanding these imbalances is crucial to ensuring the reliability and generalizability of deep learning models trained on this data.

Figure 1. Distribution of Classes in the Dataset
The data were partitioned into three subsets: training data (consisting of 499 images), validation data (125 images), and test data (156 images). This structured division aids in systematically training, tuning, and evaluating the deep learning models. The data were meticulously collected by Walid Al-Dhabyani et al. (2020) using the LOGIQ E9 ultrasound and LOGIQ E9 Agile ultrasound system at the Baheya Hospital for Early Detection & Treatment of Women’s Cancer, Cairo, Egypt. The primary motivation behind this collection was to offer a rich dataset for individuals and researchers keen on delving into the realms of deep learning applications in medical imaging.

Data Preprocessing

Before the data were introduced to our model, the following extensive preprocessing phase was undertaken to ensure the data’s cleanliness and applicability for the objectives of the research.

1. Rescaling: Pixel values of each image were rescaled to a range between 0 and 1, ensuring uniformity across the dataset.

2. Rotation: Images were randomly rotated by up to 20 degrees in any direction. This step prevents the model from becoming overly sensitive to the orientation of input scans, accommodating potential variations in future datasets.

3. Positional augmentation: The heights and widths of the images were randomly adjusted. This process captures those cases where the growth could be off-center, ensuring that the model is robust to all positional variations.

4. Shear transformation: The images underwent shearing to recognize objects even if slightly distorted. This step mitigates issues in which scans may be distorted because of operational errors, whether human-induced or machine-related.

5. Flipping: Images were subjected to horizontal flips, enhancing the model’s ability to discern objects regardless of their horizontal orientation.

6. Fill function: To ensure the integrity of images post-transformation, a fill function filled in any missing pixels that might have resulted from the rotation or flipping process.

These augmentations weren’t applied statically but were instead applied dynamically and randomized during each epoch. Rather than creating a fixed set of augmented images at the outset, our approach involved the random application of these augmentations every time an image was processed during training.

After processing the images, the model underwent training for 50 epochs with batches of 32 images. Each epoch exposed the model to 50 uniquely augmented
images, cumulatively resulting in 2,500 distinct images over the training phase. To elucidate, the batch size indicates the number of training instances used in one iteration. For our training, sets of 32 images were randomly picked, processed, and then used to adjust the model’s weights through the backpropagation algorithm. The choice of 32 as the batch size struck a balance between computational efficiency and data clarity. Additionally, the induced variability in the weight updates, inherent with smaller batches, acted as a regularization method, potentially assisting in avoiding local minima and promoting a more generalized model.

The entire dataset was apportioned into training, validation, and test subsets, as previously detailed. Postprocessing snapshots of the data are presented in Figure 2. It is noteworthy that although the training data experienced the full suite of augmentations, the validation set was solely rescaled. This distinction ensures that the validation set remains untouched by augmentations, offering a genuine performance metric. The test set remained isolated from this process and was reintroduced at the culmination of the research to assess the model’s proficiency.

Figure 2. Ultrasound Breast Cancer Scans
Model Architecture

Flowchart of Model Development

The flowchart in Figure 3 provides a succinct visual representation of the stepwise progression undertaken in our study’s model-development phase. Beginning from data collection, it charts a clear trajectory through preprocessing, model training, validation, and, ultimately, testing. This graphical representation facilitates a quick understanding of the workflow and ensures clarity in the process of our deep learning implementation for breast cancer classification.

Figure 3. Flowchart of Model Development
Transfer Learning with MobileNetV2

Transfer learning has emerged as a transformative approach to creating CNN classification models. By leveraging a pretrained model, we can expedite the learning process and subsequently refine it to the specificity of our dataset. The architecture selected for this endeavor was MobileNetV2.

Developed by Google researchers, MobileNetV2 is renowned for its efficiency, making it apt for mobile and edge devices. Trained on the expansive ImageNet dataset, MobileNetV2’s convolutional base is equipped with a nuanced understanding of myriad image categories. This pretraining offers an invaluable starting point. Given the specificities of our dataset vis-à-vis ImageNet’s multiclass nature, however, certain adaptations were necessary.

To bridge the gap between ImageNet classes and our data categories, the top layers of the MobileNetV2 model, which are more task-specific, were excluded to ensure they didn’t impose any biases from the pretrained weights (Gulzar, 2023). The architecture was then fine-tuned to accommodate our dataset with an input shape of (150, 150, 3), aligning it with the unique morphology of our ultrasound image scans.

After the model’s tailoring, an important step was to freeze its initial layers. Convolutional layers in architectures like MobileNetV2 present a hierarchical stratification of visual features. Initial layers predominantly discern basic image constructs like edges and textures. As one progresses deeper into the network, layers discern more intricate, dataset-specific features. By deploying transfer learning, it is prudent to freeze the initial layers. This retains the generalizable features learned from the ImageNet dataset, ensuring that they don’t unduly influence the model’s fine-tuning on our specialized dataset. This strategy ensures a harmonious blend of general visual understanding and specific feature recognition, pivotal for our research goals.

Hyperparameter Tuning and Calibrating Deeper Layers with MobileNetV2

Fine-tuning models, especially when leveraging the transfer learning approach, ensures that the trained architecture adapts astutely to the unique nuances of the project-specific dataset rather than merely inheriting patterns from a generic source like ImageNet. This calibration is vital for precise interpretation, especially when transitioning datasets.

Regarding MobileNetV2’s architecture, specific nuances were attended to during this research. Batch normalization layers, though often inherent in many deep learning models, were purposely excluded from this iteration. This was done to avoid distribution statistic discrepancies. In essence, pretrained models bring along with them specific statistics—the mean and variance from their originating dataset. Using these exact metrics could muddle training when applied to a divergent dataset.
Certain layers intrinsic to MobileNetV2 also underwent changes. While MobileNetV2 is designed for efficiency with depth-wise separable convolutions, our task required some tweaks for optimal results. By bypassing some of MobileNetV2’s traditional configurations, the model was augmented with a flattened layer, making the output from the previous layers amenable to a dense structure. This flattened layer was succeeded by a dense layer containing 1024 nodes, a dropout layer to counter overfitting, and another dense layer with three nodes, aligned with our project’s categorical outcomes.

The “flatten” layer deserves emphasis. It transitions feature maps from the preceding pooling layer into a singular dimensional vector. Such an alteration is pivotal, bridging the convolutional output to the dense layers. The introduced 50% dropout layer wasn’t a capricious decision either. It was interwoven to mitigate overfitting and to ensure a balanced neuron interplay, preventing any undue dependencies and thereby facilitating balanced data processing.

Complementing manual calibrations, the study made use of the Keras Tuner for a more systematic hyperparameter optimization. A popular choice for fine-tuning CNN models, the Keras Tuner discerns optimal configurations, from learning rates and dropout percentages to node quantities in dense layers and the most congruous activation functions (Saleh et al., 2022). Such a meticulous endeavor ensures the MobileNetV2 model’s peak performance for the ultrasound images under scrutiny.

**Evaluation Metrics for Medical Diagnosis**

The rigor of evaluating the performance of deep learning models in medical diagnostics is unparalleled. A misclassification not only bears direct consequences but also stresses the need for transparent and comprehensive metrics. To this end, an array of metrics was diligently selected to evaluate our MobileNetV2 model tailored for tumor image classification.

**Classification Accuracy and Precision**

Accuracy represents the proportion of correct predictions to the total predictions. It is vital in medical imaging as a basic measure of a model’s capability to correctly identify tumor classifications.

\[
\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}
\]

The value of precision, particularly in medical diagnostics, in which false positives can lead to unnecessary treatments or interventions, cannot be overstated. Precision measures the consistency of the model’s outcome for a specific classification.
Loss Value

Especially crucial in multiclass classification tasks in medical imaging, where errors can have grave implications, the categorical cross-entropy loss value function was employed to direct weight adjustments during the training phase, bridging the gap between predicted and actual values.

AUC

AUC, as related to the ROC curve, gauges model performance over various threshold values (Vakili et al., 2020). In diagnostics, in which the binary distinction between malignant and benign tumors is pivotal, AUC’s role is magnified. The computation considers the balance between the true positive rate (recall) and the false positive rate:

\[
\text{True Positive Rate} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \\
\text{False Positive Rate} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}}
\]

Matthews Correlation Coefficient

MCC is an integral metric when dealing with datasets showcasing uneven class distributions, often a reality in medical datasets where certain conditions are rare (Vakili et al., 2020). MCC’s formula encompasses all facets of the confusion matrix:

\[
\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP) \times (TP+FN) \times (TN+FP) \times (TN+FN)}}
\]

Precision-Recall Area Under Curve (PR AUC)

Vital for cases in which imbalances in data are the norm or when precise detection (as with rare diseases) is the focus, PR AUC embodies a model’s finesse in harmonizing precision and recall across assorted thresholds.

By incorporating these metrics into the evaluation phase, the study seeks to guarantee not only the model’s precision but also its reliability and reproducibility, especially in the high-stakes arena of medical diagnostics. Continuous monitoring of these metrics throughout the training and validation stages facilitated the progressive refinement of the MobileNetV2 model, ensuring its readiness for pivotal medical applications.
Results and Discussion

This section offers an in-depth comprehensive analysis of the CNN model's overall performance within the area of breast ultrasound class. Our study remains intently targeted at achieving the vital objective of correctly discerning normal, benign, and malignant ultrasound pictures using the model. In this section, we delve into the evaluation of the model's efficacy in achieving the objective of our study. Figure 4 provides all essential model evaluation metrics interpreted in this section.

Figure 4. Essential Model Evaluation Metrics

Accuracy

The derived accuracy metric, quantified at 0.82, stands as a testimony to the model’s prowess in correctly classifying approximately 82% of the analyzed ultrasound samples. This metric holds profound implications in the context of tumor classification, and the capability to reap such an accuracy is pivotal, as it substantiates the model’s ability as a reliable tool in the domain of medical imaging and diagnosis. This degree of precision is crucial for supporting clinical judgment and therapeutic direction, especially when it comes to spotting possible cancers. To ensure early diagnosis and proper patient management, the model’s ability to provide accurate classification is crucial.

Precision

The model has the ability to correctly identify 83% of anticipated positive samples, as indicated by a precision of 0.83 shown in Figure 4. This precision metric
highlights the model’s effectiveness in lowering the incidence of false alarms. Such precision has significant ramifications, especially in the delicate area of medical diagnosis, reducing unnecessary procedures and consequent patient emotional discomfort.

Recall

The recall metric (0.81) demonstrates the model’s exceptional ability to accurately identify and capture 81% of positive cases in the dataset. This proficiency in minimizing false negatives is crucial for medical diagnostics, as missing true positive cases can have significant consequences. This high-recall data point enables the timely detection of potential malignancies or abnormalities in ultrasound images, improving patient safety and medical decision-making. The model’s potential in medical practice, tumor classification using deep learning, and healthcare standards enhancement further emphasizes its value.

ROC-AUC

The ROC-AUC score of 0.94 indicates excellent model performance, with a curve area approximating 1. The ROC curve represents the model’s true positive rate against its false positive rate and the model’s discriminatory power, whereas the AUC measures the model’s effectiveness in distinguishing positive and negative cases. The high ROC-AUC score underscores the model’s proficiency in tumor classification, highlighting its potential clinical utility. Figure 5 showcases the model’s specialized discriminative capabilities. A detailed examination of the ROC scores across each class provides critical insights into the model’s distinguishing power. Notably, class 0 performance shines distinctly, registering a remarkable ROC score of 0.97. This score signifies the model’s superior aptitude to discern instances of class 0 from the broader dataset. Such a high ROC score elucidates the model’s competency in curtailing false positives while concurrently maximizing the true-positive rate. This ability is paramount in the realm of medical diagnostics.
The PR AUC score of 0.88 indicates the model’s performance across precision and recall thresholds, providing valuable insight into imbalance datasets or clinically significant classes. It evaluates the model’s ability to balance precision (identifying positive cases) and recall (capturing all positive cases). This high PR AUC score enhances the model’s suitability for real-world clinical applications and strengthens its value in deep learning–based tumor classification.

The PR curves for each class in Figure 6 indicate AUC values of 0.86, 0.88, and 0.91 for classes 0, 1, and 2, respectively. These curves exemplify the intricate equilibrium between accurately identifying positive instances and minimizing false positives.
The PR curves offer an intuitive representation of the interplay between precision and recall across diverse thresholds. The associated AUC values capture this balance statistically, aiding in the selection of optimal decision thresholds aligned with specific clinical requirements.

The model’s PR dynamics are further illustrated by the PR AUC scores for each class: 0.86 for class 0, 0.88 for class 1, and 0.91 for class 2. These scores adeptly highlight the model’s refined performance in achieving the best true positive detection while judiciously limiting false positives.

**MCC**

The MCC score of 0.74 indicates the model’s strong correlation between observed and predicted classifications, encompassing true positive, true negative, false positive, and false negative cases. This metric evaluates the model’s performance across all four outcomes in medical diagnostics, where sensitivity and specificity are crucial. A high MCC indicates a stronger agreement between predictions and the ground truth, indicating consistent and reliable results (Chicco & Jurman, 2020). This score demonstrates the contribution opportunity to accurately diagnose while simultaneously aligning with medical practice demands, ultimately validating its role in advancing deep learning–based tumor analysis.

This score is the lowest because the MCC regards the negative class samples (true negative and false negative) highly. This highlights that the model is not always accurate when discerning when someone does not have breast cancer, leading to an
increase in claims of malignant tumors when the tumors are actually benign or normal.

Model Performance Analysis

Validation Analysis

To comprehensively evaluate the performance of our tumor classification model, we conducted a detailed analysis of key performance metrics across training epochs. A constant rising trend is shown in evolution of validation accuracy over epochs in the Validation Accuracy graph of Figure 7. This improvement demonstrates the model’s ability to recognize complex patterns in breast ultrasound pictures and to modify internal representations to produce precise predictions. Figure 7 also displays the validation precision over epochs.

It is noteworthy that accuracy shows an initial rise before leveling off in later epochs. This finding highlights the model’s capability to recognize positive examples with accuracy, but it also raises the possibility of a saturation point for precision advances. The validation recall graph in Figure 7 shows a steady and significant increase. The model’s ability to accurately identify real positive cases is demonstrated by its high recall values, which is an important characteristic in medical applications where minimizing false negatives is crucial. Lastly, Figure 7 also displays a validation loss graph. The model’s increased convergence and learning are indicated by the observed consistent decline in loss values. This decrease in loss is consistent with the rising trends shown in recall, accuracy, and precision, which supports the model’s overall effectiveness.

Figure 7. Model Performance Analysis Graphs
Confusion Matrix

A confusion matrix is a crucial tool for evaluating the model’s classification performance, succinctly providing a thorough picture of the anticipated versus real labels, enabling a straightforward assessment of the model’s advantages and weaknesses. The diagonals of the matrix stand in for accurate classifications, and certain entries show occurrences of true positives, true negatives, false positives, and false negatives.

The diagonals (35, 34, 37) for the described matrix (35, 1, 1; 6, 34, 9; 2, 4, 37) show precise classifications for classes 0, 1, and 2 (Figure 8). The performance of the model can be improved by highlighting probable areas for improvement, such as the nine occasions in which class 1 was incorrectly categorized as class 2. The need to emphasize how crucial it is to recognize classification errors is especially clear in the field of medicine, where incorrect classification can have serious repercussions. Inaccurate classifications may result in postponed or ineffective medical measures, endangering the health and outcomes of patients.
Image Intensity Analysis

The distribution of image intensities (Figure 9), when examined across all categories, illustrates a clear right-skewed pattern. This observation suggests that most images primarily feature lower-intensity values, witnessing a gradual reduction in frequency as we move toward the higher-intensity spectrum. Such a distributional characteristic implies that the dataset’s mean is inclined to be greater than its median, which the boxplot statistics further corroborate.

Figure 9. Distribution of Image Intensities

Figure 10. Boxplot Distribution of Image Intensities by Class
Focusing on the Normal Class image intensities in Figure 10, we see they range between 0.16 and 0.44. The central spread of data, representing 50% of these images, resides within the 0.25–0.34 range, emphasizing a consistent luminosity prevalent in normal images. The median intensity value is recorded at 0.30, underpinning this observation.

Benign Class images exhibit a broader spectrum of intensities, spanning from 0.16 to 0.51. The interquartile range, indicative of the central half of the data, stretches between 0.28 and 0.41. This expanded range, juxtaposed against the normal images, points toward an augmented variability in the illumination characteristics of benign images. Their median intensity, pegged at 0.33, is slightly elevated when compared to the Normal Class.

Malignant Class images, intriguingly, present the widest span of intensities, which are contained between 0.14 and 0.47. The middle 50% of the data is ensconced between 0.22 and 0.36. The median intensity value for malignant images is charted at 0.28, indicating a tendency for these images to possess lower intensities in comparison to Benign and Normal Class images. Such trends might be emblematic of the inherent attributes or behavior of malignant tissues under imaging.

Across all categories, there is a pronounced absence of outliers, highlighting consistent imaging and processing standards in the dataset. This consistency is particularly vital in the domain of medical imaging, ensuring diagnostic dependability. Although the intensity spans for individual categories exhibit overlapping tendencies, their central tendencies and dispersion metrics are distinct. Such nuances suggest that image intensity may not be the sole determinant for classifying an image, but it undeniably plays a significant role in assisting such categorization. Our analytical journey into the realm of image intensities yields crucial insights into the distinguishing features inherent to images of the normal, benign, and malignant classes. The observed disparities in the intensity profiles of these categories suggest the potential influence of either intrinsic tissue properties or imaging techniques in determining the resultant image outcomes. These findings, when harnessed effectively, could significantly elevate the precision of medical diagnostics.

Error Analysis

In our quest to further refine and optimize our model, conducting a thorough error analysis offers an invaluable perspective. Such an analysis provides insights into where our model is faltering, allowing us to implement precise solutions.

Analyzing Misclassified Images

Upon reviewing the table of misclassified images (Figure 11), it becomes evident that certain images are consistently misinterpreted by our model. This table
depicts side-by-side comparisons, showcasing the true label of an image alongside the predicted label by our model. For instance, consider an image within this table: The true label might be *benign*, but our model erroneously predicts it as *malignant*. On examining this image, we might observe that its characteristics are borderline, making it challenging even for a human expert to classify. Such a subtle differentiation point can throw off our model, causing these misclassifications. This form of visualization helps us understand the specific instances and scenarios where our model is underperforming. By taking a closer look at these misclassified images, we can infer potential reasons for these errors and strategize how best to rectify them.

**Potential Reasons for Misclassification**

*Data Quality*

One of the fundamental pillars of any machine learning model is the quality of its training data. It is imperative to ensure that our dataset is free from label noise. Mislabeled images can cause the model to learn incorrect patterns, thereby affecting its generalization capabilities on unseen data. Regular audits and checks on the dataset can help in identifying and rectifying such issues.

*Data Representation*

Some images might inherently possess a high degree of complexity, with subtle features playing pivotal roles in classification. If our current model architecture isn’t attuned to capture these nuances, the result could be misclassifications. This calls for a reevaluation of our model’s architecture and perhaps a move toward a more sophisticated design.

*Lack of Relevant Data*

A balanced and diverse dataset is a cornerstone for robust model performance. If our dataset lacks ample representation for certain classes, the model’s performance for those classes would invariably suffer. Though techniques like data augmentation offer some respite, there is no substitute for acquiring more real-world, varied samples for underrepresented classes.

*Complexity of the Model*

Although simplicity has its merits, an overly simplistic model might not do justice to a complex dataset. It might be worth experimenting by ramping up the model’s complexity, incorporating more layers, or even integrating a different pretrained model.
Hyperparameters

Every model has a set of hyperparameters that allows it to perform optimally. Our current choice of hyperparameters such as regularization rates, learning rates, and batch sizes might not be in this optimal range. Engaging in systematic hyperparameter tuning can unearth a combination that boosts model performance.

Class Imbalance

In scenarios in which certain classes overshadow others in the dataset, there is a tangible risk of the model developing a bias. This bias can skew predictions in favor of the dominant classes. Addressing class imbalances through techniques such as oversampling, undersampling, and application of class weights can pave the way for a more balanced and objective model.

Although our model demonstrates commendable performance, there is always room for refinement. Through systematic error analysis and addressing the potential reasons for misclassifications, we can inch closer to a model that not only performs better but also resonates more with real-world complexities.

Figure 11. Misclassification Diagram
Comparative Analysis of our Findings with Contemporary Research

Mishra et al.’s Radiomics-Based Approach vs. MobileNetV2-Based Classifier

Using the same dataset to tackle breast ultrasound tumor classification, Arnab Mishra and colleagues (2021) applied a radiomics-based approach. Though similar to the research done above, the study by Mishra et al. approached classification by focusing heavily on extracting various image features from the region of interest and incorporated a recursive feature-elimination method for selecting the most pivotal features. In contrast to this approach, the MobileNetV2 model utilized for this study zeroes in on the transfer learning paradigm, specifically leveraging a pretrained model and meticulously fine-tuning it to harmonize with the dataset given.

A pivotal aspect that is pertinent to this study is the inherent class imbalance given in the dataset. To address this, Mishra et al. (2021) employed synthetic minority oversampling (SMOTE), whereas our study simply used random oversampling because of computational restraints. In terms of feature significance, Mishra et al.’s study revealed that the shape, texture, and histogram-oriented gradient features stood out as paramount for the classification endeavor; however, the MobileNetV2 methodology bypasses manual feature extraction by capitalizing on pretrained weights inherited from the ImageNet dataset.

The approach of Mishra et al. (2021) is radiomics-centric, emphasizing the extraction of a diverse range of image features from the region of interests and further refining them using recursive feature elimination. Their research achieved an impressive accuracy of 0.974, an F1 score of 0.94, and an MCC value of 0.959 on the BUSI dataset. In stark contrast, the MobileNetV2-based classification, which utilizes the transfer learning paradigm and fine-tunes the architecture for the specific dataset, secured an accuracy of 0.82, a precision of 0.83, and a recall of 0.81. Additionally, while the ROC-AUC for the MobileNetV2 method was 0.94, almost mirroring that of Mishra et al., the MCC was slightly lower, at 0.74, indicating a lesser degree of reliability in the binary classifications made by the model.

Considering practical applicability, the stellar outcomes of Mishra and colleagues (2021) indicate a strong potential for real-world medical diagnostic application. Simultaneously, the MobileNetV2 methodology, despite a slightly lower accuracy, suggests immense promise for real-time applications, especially on resource-constrained platforms such as mobile devices, prioritizing rapid responses with substantial precision.

Cruz-Ramos et al.’s Hybrid Feature Fusion vs. MobileNetV2-Based Classifier

Cruz-Ramos et al. (2023) ventured into the realm of breast tumor classification by employing a hybrid methodology that fused deep learning features with traditional
handcrafted ones. Their computer-aided diagnostic (CAD) system’s core revolved around the amalgamation of DenseNet 201 architecture and traditional handcrafted features, including histogram of oriented gradients (HOG), ULBP, perimeter area, area, eccentricity, and circularity. The fusion process was reinforced using genetic algorithms and mutual information selection, followed by employing classifiers such as XGBoost, AdaBoost, and multilayer perceptron (MLP).

Their study spanned two imaging modalities, mammography and ultrasound, leveraging datasets mini-DDSM and BUSI. Their results were indeed impressive, with accuracy of 97.6%, precision of 98%, recall of 98%, F1 score of 98%, and IBA of 95% on the aforementioned datasets. This holistic approach speaks volumes about the potential of combining traditional image-processing techniques with modern deep learning architectures to obtain high-precision diagnostic results.

In contrast, the research based on the MobileNetV2 architecture concentrated solely on deep learning, adopting a transfer learning paradigm and aligning the model with the specificities of the BUSI dataset. Results obtained from this method exhibited accuracy of 0.82, precision of 0.83, recall of 0.81, ROC-AUC of 0.94, and MCC of 0.74. While both methods achieved commendable results, Cruz-Ramos et al.’s (2023) hybrid method, which combined deep and handcrafted features, exhibited superior performance metrics.

The implications of these findings suggest that while deep learning approaches like MobileNetV2 offer substantial promise, there might be untapped potential in merging traditional image-processing techniques with modern neural network architectures for enhanced diagnostic precision in the domain of breast tumor classification.

Labcharoenwongs et al.’s Tumor Detection with YOLOv7 vs. MobileNetV2-Based Classifier

Labcharoenwongs and team (2023) delved deep into breast tumor detection, classification, and volume estimation using a comprehensive deep learning approach. Their primary objective was to facilitate radiologists with automated tools that can complement their decision-making, especially when manual analysis is influenced by radiologist skill levels and image quality. For this, they utilized the You Only Look Once version 7 (YOLOv7) architecture to detect, localize, and classify tumors from ultrasound images. Their dataset comprised 655 images, with a mix of benign and malignant samples, and they augmented this dataset with various methods such as blurring, flipping, and introducing noise.

Their results are worth noting. The YOLOv7 architecture achieved a confidence score of 0.95 for tumor detection. In terms of lesion classification, their
model achieved accuracy of 95.07%, sensitivity of 94.97%, specificity of 95.24%, positive predictive value of 97.42%, and negative predictive value of 90.91%.

When this is juxtaposed with the research deploying the MobileNetV2 architecture, the following observations arise: The MobileNetV2 model exhibited accuracy of 0.82, precision of 0.83, recall of 0.81, ROC-AUC of 0.94, and MCC of 0.74. Labcharoenwongs et al.‘s (2023) model, driven by the YOLOv7 architecture, seems to achieve higher accuracy, sensitivity, and specificity. This suggests that although the MobileNetV2 architecture is adept at handling breast ultrasound image data, architectures such as YOLOv7, tailored more toward real-time object detection, might offer an edge in terms of tumor detection and classification in this specific context. The potential integration of such systems into conventional ultrasound machines, as proposed by Labcharoenwongs and team, underscores the evolving landscape of medical diagnostics and the role of deep learning therein.

Deployment

The deployment phase marks a pivotal transition in a machine learning project, from the confines of research and development to tangible, real-world applications. One of the chosen platforms for such transitions, especially in the realm of data-driven insights, is Streamlit. Recognized within the data science ecosystem, Streamlit is specialized in catalyzing the creation of interactive web applications. Its distinct advantages, such as enhanced interactivity and swift deployment capabilities, position it as a preferred choice for introducing advanced algorithms to diverse user groups. The adoption of Streamlit for this project was informed by Streamlit’s robust scalability, which can accommodate the intricacies of complex models, combined with an interface that does not require exhaustive web-development expertise.

UI/UX

At the user-interaction frontier, the Streamlit interface has been designed to maximize both functionality and user-friendliness. The interface facilitates the direct upload of breast cancer images, supporting a gamut of formats, including PNG, JPG, and others. Upon the initiation of an upload, the application provides immediate feedback, thereby negating ambiguities. Moreover, real-time visualizations of the uploaded images enhance user interaction, with features such as zoom-in for detailed insights. Predictions, central to the application, are rendered transparently, frequently in percentage form. These values, indicative of malignancy likelihood, are reinforced with color-coded distinctions for immediate comprehension.
Back-end Operations

While the front-end addresses user interactions, substantial computational operations underlie the application. To prioritize user privacy, the system is designed to process images without long-term retention. The preprocessing mechanisms are streamlined to echo the research phase methodologies, ensuring a high degree of accuracy. Response times for predictions have been optimized to be nearly instantaneous, facilitating user engagement.

Code Implementation

The deployment code commences with necessary imports, notably Streamlit, TensorFlow, and PIL, ensuring the required libraries are available. Once initialized, a pretrained model, optimized for tumor classification, is loaded into the application. A function, predict_tumor(), is defined to resize, preprocess, and feed the image to the model, returning the predicted class of the tumor. Enhancing the user interface, a centralized header introduces the application, followed by prompts guiding the user to upload an image. Upon image upload, the application displays the chosen image, processes it, and presents the prediction. This deployment, while providing an immediate application of machine learning research, also lays the foundation for further refinement and expansion. Feedback loops, system integrations, and richer datasets stand as promising avenues for future iterations. Figure 12 provides an image of our simple preliminary deployment of the created model.
Limitations

Model development is heavily affected by the limitations imposed by hardware and software complexities. The implementation of larger and more-complex models can be restricted by a lack of computational resources, which could impede the investigation of cutting-edge architectures. Additionally, the use of intensive computational methods without necessary computational resources would limit the depth and complexity that the model’s design could support. The time commitment needed to train models, especially when dealing with large datasets, is a significant obstacle. The rapid iterative creation of models is constrained by training times, which
can grow significantly when dealing with a large dataset. Long training times further hinder the model’s instant applicability when switching to real-time applications, making it less suited for time-sensitive scenarios like rapid diagnosis in healthcare settings. Another significant obstacle is the complex interplay between model complexity and memory allocation. The maximum batch sizes that can be utilized during training are constrained by memory capacity, which may reduce the model’s ability to recognize complicated patterns in the data. These memory constraints could be made worse by complex models with many layers and parameters, which could exceed the amount of memory that is available. This can reduce the model’s ability to capture subtle details, which will affect the accuracy of its classification and its ability to discriminate.

Dataset-Derived Limitations

The dataset’s intrinsic biases have the potential to majorly influence the effectiveness and generalizability of the model. These biases may result from a variety of causes, including demographic, geographic, or historical. Biases may develop as a result of differences in patient demographics, scanning technologies, or data-collection techniques across various geographic or temporal contexts. Such biases can have two negative effects: they can result in an unfair depiction of some classes, which distorts the model’s impression of their predominance, and they can create an unnatural bias in favor of dominating classes.

A special set of difficulties is presented by the authenticity of medical images, which is essential for reliable diagnosis. Resolution, contrast, and noise variations in an image’s quality can lead to complexities that prevent accurate classification. The model may lead to misdiagnosis due to poor image quality. As a result, misclassification is more likely to occur and the system is less resilient in real-world situations when image quality may be reduced.

Class-distribution imbalances can have a major effect on how well a model performs after being trained. Situations in which particular classes are overrepresented constitute a serious problem. The model may develop a preference for overrepresented classes as a result of training on unbalanced datasets, potentially pushing minority classes to the edge of their discriminatory power. Additionally, the risk of misclassification in crucial medical contexts may increase as rare classes appear because of their dominating counterparts.

Model Generalizability

The imperative evaluation of the model’s behavior under scenarios of data variability beyond its training ambit assumes paramount significance. The model’s performance may be disrupted when exposed to data distributions, imaging
techniques, or clinical settings that differ significantly from its training environment. Its ability to identify strong features and patterns that go beyond the limitations of mixed datasets is what gives it the ability to extrapolate knowledge across diverse fields.

The flexibility of the model to adapt to unknown instances of unknown data is a need for the model to perform well. While some models exhibit remarkable transferability, others could require recurring retraining to accommodate the changing data landscape. The creation of systems for constant monitoring, recalibration, and harmonious integration within the changing therapeutic paradigm is also necessitated by the specter of obsolescence.

The empirical testbed of the model, set amongst the complicated reality of real-world circumstances, serves as the model's crucible for validation. The gap between theoretical power and practical application is highlighted by the disjunction in performance indicators when the model moves from the organized confines of controlled experiments to the nuanced details of real clinical usage. These inequalities highlight the importance of real-world evaluations, which provide essential information that informs model modifications and aligns the model's predictive power with the complex and ever-evolving landscape of clinical treatment.

Deployment Issues

The model's capacity to adapt to a growing user base and expanding data streams will determine how well it is implemented. As demand grows, the difficulties will lie in overcoming possible constraints linked to memory allocation, processing resources, and network bandwidth. To guarantee the model is capable of scaling, it is crucial to implement strong parallelization algorithms, load balancing systems, and use distributed computing paradigms. Additionally, effective management of resources is required, which may be addressed by designing an efficient architecture using microservices or containerization.

The model's performance is assessed in the setting of crucial medical decision-making, based on its capacity to offer real-time insights, combining diagnostic precision with prompt response. To achieve real-time processing, careful optimization work must be done to reduce latency and processing delays. This necessitates thoughtful model complexity management, deliberate algorithmic optimization, and efficient feature extraction. It is crucial to strike a balance between model sophistication and computing efficiency to allow real-time functionality without sacrificing diagnostic efficacy.

Complex issues with interoperability and data management arise when the model is integrated into elaborate medical systems and databases. Version-control processes, strong data pipelines, and rigorous data mapping are required to address
compatibility difficulties and ensure data synchronization. Because of differences in data formats and technologies, compatibility problems could occur, making middleware development necessary to ensure smooth communication. To provide a dependable exchange, efficient data transport requires strong compression methods and low latency protocols. Bidirectional data flow and versioning solutions are required to maintain coherence between the model and current systems because of synchronization complexity. Technical expertise and rigorous testing are vital for a successful deployment and for clinical decision-making.

A thorough evaluation of the model’s performance has revealed its advantages and drawn attention to any weaknesses that need to be considered when interpreting results or considering its use in the real world. The difficulties mentioned—which include resource restrictions, limitations arising from datasets, and model generalizability—underline the necessity of taking a cautious and knowledgeable approach. Overcoming the stated constraints is of utmost significance for future studies. Resource limitations can be eased by investigating sophisticated parallel processing architectures and optimized algorithmic implementations. Model generalizability could be improved by using transfer-learning approaches specifically designed for medical imaging data. For addressing class imbalances and improving overall performance, techniques such as data synthesis and augmentation are crucial. Including clear AI approaches could improve the clinical interpretability of data. Research frontiers must be advanced by collaborative projects integrating machine-learning experts, medical professionals, and subject specialists. By integrating these efforts, the field may overcome obstacles, cultivate reliable models ready for real-world use in healthcare settings, and ultimately lead to enhanced diagnostic accuracy and patient outcomes.

Conclusion

This investigation into breast cancer classification culminates in a model underpinned by machine learning methodologies, demonstrating a profound synergy between oncological diagnostics and computational techniques.

Research Contributions and Discoveries

This study aimed to design an efficient deep-learning model for the classification of breast cancer images. Through employment of CNNs and optimization of various hyperparameters, a discriminative model was successfully trained to distinguish between normal, benign, and malignant tumor presentations. Key insights into the distribution of image intensities and their relation to tumor classifications were discerned. The deployment of the model into a Streamlit web
application exemplifies how academic research can be translated into tangible, user-centric tools.

Implications for Healthcare

The implications of this model are not limited to the domain of machine learning; they resonate profoundly within healthcare. With this enhanced diagnostic aid, clinicians can achieve quicker, more accurate diagnoses, expediting patient treatment plans. By automating a segment of the diagnostic process, the model facilitates optimal resource allocation in healthcare settings, fostering both efficiency and precision.

Future Directions

The model’s current performance, while commendable, provides a canvas for several future trajectories:

*Model Optimization*

Continuous feedback from the deployed tool can help in gradient descent-based refinements, iteratively reducing loss and improving the model’s precision and recall metrics.

*Diverse Architectures*

Although the current model leverages a particular CNN architecture, exploring alternative configurations, such as residual networks (ResNets) or transformer-based models, could further elevate its performance.

*Expansion of Datasets*

One of deep learning’s core tenets is its capacity to thrive on vast datasets. Incorporating larger and more diverse datasets can enhance the model’s generalization, reducing overfitting and bolstering its robustness in varied diagnostic scenarios.

*Addition of Further Classes*

It is evident that the dataset currently used is limited in the number of classifications it provides. A big step in this research would be to gain data that provide deeper classifications—for example, delving more deeply into types of malignancies (e.g., adenocarcinoma, melanoma) and types of benign tumors (e.g., lipomas, adenomas). Another important aspect to note is datasets that contain unclassified
images, or images that would be hard for even humans to identify. Having these broader class definitions will significantly increase the robustness of the deployment.

In summary, though it may be seen that this endeavor did not maintain levels of accuracy and other metrics, this endeavor underscores the promise and potential of deep learning in oncological diagnostics. As the confluence of machine learning and healthcare continues to expand, there is great anticipation for even more refined tools and methodologies that can further the mission of early and accurate cancer detection.
References


OPENING AN EXISTENTIAL WINDOW INTO WUTHERING HEIGHTS AND THE METAMORPHOSIS

SARAH H. LOGNION, McNEESE STATE UNIVERSITY
MENTOR: JOSEPH TUMINELLO

Abstract

This paper explores the existential theme of authenticity in two literary works, Emily Brontë’s Wuthering Heights and Franz Kafka’s The Metamorphosis, through the lens of Albert Camus’s philosophy as presented in The Myth of Sisyphus. The analysis focuses on the characters’ struggles to carve out meaningful existences, the symbolism of closed-in structures and the absurd, and failed attempts at rationalization. Through a literary exploration, this paper aims to acknowledge existential dilemmas presented in Brontë’s and Kafka’s works and the importance of recognizing the absurdity of life.

Albert Camus, a prominent twentieth-century existential philosopher, wrote about Franz Kafka’s literature, “The whole art of Kafka consists in forcing the reader to reread. His endings, or rather his absence of endings, suggest explanations which, however, are not revealed in clear language. . . . Sometimes there is a double possibility of interpretation, whence appears the necessity for two readings.”¹ Pauline Nestor wrote in review of Emily Brontë’s Wuthering Heights that it “[failed] to conform to my expectations.”² Of course, Nestor’s expectations of the novel had been shaped by romantic film adaptations and the fact that she first read the book at the age of twelve, but the point remains the same. Brontë’s and Kafka’s literature are the type that people tend to misinterpret because of the subtle symbolism that is elusive in a first reading. When one typically thinks of Emily Brontë’s novel Wuthering Heights, existentialism might not be the first topic to come to mind. Brontë is not known for being a writer within the existential movement, but the existential angst presented in Wuthering Heights can be compared to topics explored in Camus’s The Myth of Sisyphus and Franz Kafka’s complex novella The Metamorphosis.

Both Wuthering Heights and The Metamorphosis are concerned with characters who have failed to carve out meaningful existences for themselves because of both their own shortcomings and the circumstances in which they find themselves. Both works demonstrate the perils of inauthenticity, since most of the internal strife in the characters’ lives is caused by a lack of authenticity. The characters, Catherine Earnshaw in Wuthering Heights and Gregor Samsa in The Metamorphosis, are victims of the absurd rather than finding within it as Camus states “my revolt, my freedom, and
my passion.” By examining *Wuthering Heights* and *The Metamorphosis* within the context of Camus’s *The Myth of Sisyphus*, the reader can finally delve into that second reading of existential symbolism.

**Doors and Windows**

Albert Camus writes of the absurd walls found in our worlds, “At the heart of all beauty lies something inhuman…. The world evades us because it becomes itself…. It withdraws at a distance from us.” This is the open secret of the absurd, or the sense that human existence in the universe is meaningless. The world itself is not absurd, nor is life itself, but the two together create the absurd. When a person has a glimpse of the absurdity of life, the world becomes inhuman and distant as the realization sets in that the world is not a beautiful place according to human standards. Camus’s response to the absurd is to rebel against it with passion by creating meaning in life and thus attempting to find the self. If one does not revolt, Camus surmises, the individual will be left to live a closed-off existence by lacking the creation of any essence in life. The individual will either go through life inauthentic or will no longer be able to tolerate the absurd and will respond by suicide. Although no character in *Wuthering Heights* or *The Metamorphosis* chooses suicide, the topic is mentioned a few times in *Wuthering Heights*, particularly in reference to Catherine, Heathcliff, and Hindley. The characters in both works are left in a closed-in state that prevents both their intentional suicides and their revolts to discover authenticity.

One main symbol found in both *Wuthering Heights* and *The Metamorphosis* is closed-in structures used to emphasize how inauthenticity leaves an individual cut off from the self. Instances of inauthenticity will be described more in depth in other sections of the paper. The characters in both works are destined to live within trapped worlds, whether an insect’s body or an isolated house on the moor. The symbolism of the closed-in structure is presented both figuratively and literally in *Wuthering Heights* and *The Metamorphosis*. For example, in *Wuthering Heights*, the characters are isolated despite not being far from the crossroads and Thrushcross Grange. The nearby town of Gimmerton seems leagues away, and outsiders in the plot are rare. Catherine Earnshaw’s childhood bed is described as “a large oak case,” and later scenes involve characters locking themselves within rooms or being locked inside rooms by others. Several times, a character is locked outside of a room as well. When Catherine experiences her second illness, Edgar, Nelly, and the rest of the household staff are locked out of her room for three days.

*The Metamorphosis* also has a confined feeling, since the action takes place only within the Samsa house, particularly within Gregor’s room. Gregor is kept confined within his room following his transformation into a bug. His window to the past is the furnishings within the room, especially his prized writing desk. Gregor’s literal window from which he can see the hospital across the street is eventually
closed to him as he loses the crisp sight of a human and gains the sight of a bug. Li Anheng observes that, “even [though] the hospital is on the other side of Gregor’s house, his family never send[s] him to see the doctor.” The window is closed to him because help across the street is not available for his care or for his failing sight to glimpse. Not only that, but the adjacent hospital treats humans, not humans transformed into insects, so even if his family had taken him, more than likely, nothing could have been done to help his condition.

Windows to the past or outside are frequently shown as open, however, such as when Lockwood breaks the glass window in his dream. Following a misunderstanding that Lockwood is stealing a lantern, he is attacked by the dogs at Wuthering Heights and then is taken to Catherine’s childhood room to spend the night. He subsequently has a couple of disturbing dreams after reading Catherine’s old books, including the dream that he breaks the glass to stop a branch from knocking against the window, only to encounter the ghostly hand of Catherine as a child.

Catherine’s opening the window when under her attack of “brain fever” demonstrates a similar look to the past. Catherine opens the window and experiences hallucinations of the lights from Wuthering Heights being visible even though Nelly describes the lights as never having been visible. She also claims to talk with Heathcliff about not traveling through the Kirkyard and says, “He’d rather I come to him.” A child claims to see both Heathcliff and Catherine wandering together near Wuthering Heights following Heathcliff’s death. Nelly (the housekeeper) finds Heathcliff’s window open and his room door locked upon his death, demonstrating that his door to life and others is closed and his window to the past, and presumably Catherine, is open.

In The Metamorphosis, the Samsa family began leaving Gregor’s door open in the evenings toward the end of his life. This provides Gregor with yet another glimpse of how life at the Samsa house was forever changed and caused him to reminisce about the days before he became an insect, thus providing a gloomy view of the past. After his death, Gregor’s door and window are opened wide and he is removed by the charwoman. The family takes a holiday and leaves the confines of home and work the day of Gregor’s death. In Wuthering Heights, the closed-in structures appear less concrete around death. Catherine’s death brought new life, her daughter Cathy, into Thrushcross Grange, and Catherine and Heathcliff’s ghosts are believed to roam near Wuthering Heights. Heathcliff tries to dig up Catherine’s body following her death and sees her corpse eighteen years after her death.

An existential perspective into why closed-in structures open around death is that death is a complete mystery to the existentialist. Vilma Irén-Mihály states that, “[a]ccording to existentialists, we ourselves are the question, and therefore we cannot ever ask the question why about ourselves. Thus, existentialism highlights the importance of the subjective self, while this subjectivity becomes the source of our
Kafka and Brontë, though not specifically existential philosophers, both demonstrate belief in the idea that humans cannot answer the question “why” about themselves. Gregor can attempt to express himself all he wants by walking on the walls and ceiling of his room, Catherine can become as much or as little of a lady as she likes, and Heathcliff can dig up Catherine’s body repeatedly, yet none of these three is any closer to answering the question “why” about their existence and ending of existence (death).

Existentialism can comment upon the lack of externally imposed meaning on life and the detrimental effects of using a particular religious ideology or belief system to impose meaning but cannot state with certainty whether there is an unseen external meaning to life, or whether life continues after death. Albert Camus states of death and life, “By the mere activity of consciousness I transform into a rule of life what was an invitation to death.” Because Camus cannot peer into the abyss of death, he chooses life despite realizing that the world and his existence have no logical meaning to impart. He chooses attempting to give meaning to his subjective self instead of demanding the answer to “why” his entire life and creating undue misery. Along this line of reasoning, for Camus, one must turn away from the temptation of the open window of existential angst over questions that are not intended for humans to answer and must resort to making the best of the closed-in room, since that is the way to create meaning in life; thus, the windows must be left open following death because the existentialist’s guidance for life and the self ends without a concrete conclusion.

Life’s Wake-Up Calls

Another piece of existential symbolism featured in both Wuthering Heights and The Metamorphosis is that characters have missed their “wake-up calls” in life. Gregor literally misses his alarm clock on the morning he transforms into a bug, and he has missed his wake-up call figuratively in life as well. He labors as a traveling salesman because his family is dependent on him even though he despises the job. As Gregor puts it, “What an exhaustive profession I’ve chosen! Day in and day out on the road. Work like this is far more unsettling than business conducted at home.” Comparably, characters in Wuthering Heights have also missed their wake-up calls and go through false wake-up calls. For example, Catherine becomes close with Heathcliff as a child after initially being cruel to him. One night, after being turned out of the house by her brother Hindley and his wife Frances, Catherine and Heathcliff have their first encounter with Thrushcross Grange and its occupants. Catherine is
bitten by a dog and is left at the Grange to recover. When she returns home around Christmas, she is markedly changed for the worse, although well-bred society would have considered her improved. Hindley notes the change immediately: “Why, Cathy, you are quite a beauty! I should scarcely have known you—you look like quite a lady now.”23 But all she has done is transform into what the world and her family wanted to see, not discovered her own essence and identity. To others and initially perhaps to Catherine, it would appear she has found her essence and identity—as a fine lady, which at first pleases her. The new Catherine proceeds to scorn her childhood companion and part of herself, Heathcliff, in favor of new companions Edgar and Isabella Linton.

The backstory behind Gregor’s metamorphosis is comparable to Catherine’s. Gregor’s room is described as containing a writing desk and a picture of a lady in a gilded frame hanging on the wall.24 At one time before he sacrificed himself, Gregor had wanted to be a writer. The picture of the lady cut from a magazine may reflect Gregor’s past hopes that he could find a lady to be his wife, as well as his love of finery. But like Catherine’s transformation into a fine lady, Gregor’s transformation into the mere laboring son, brother, and employee has left him lifeless and predetermined as a paid model for a catalogue whose picture is placed in a gilded frame. Gregor is lifeless in the sense that he is no longer concerned with his own self, in favor of others. He is predetermined in that, like a model presenting clothing in a fashionable way to sell to others, he is given a set role to be the life of others and not of himself. The model is at least allowed some difference in self-expression, however, whereas Gregor lacks even this. On top of this is the gilded frame of a respectable family that Gregor has provided. Inside the gilded frame is darkness, both in the Samsa family and in Wuthering Heights, since both worlds are trapped in a vicious cycle of a lack of self-expression, rationalization, and inauthenticity.

It is intriguing that a quote from Camus’s Myth of Sisyphus echoes Gregor’s realization that he missed his “wake-up call.” Camus writes,

Rising, streetcar, four hours in the office or the factory, meal, streetcar, four hours of work, meal, sleep, and Monday Tuesday Wednesday Thursday Friday and Saturday according to the same rhythm—this path is easily followed most of the time. But one day the “why” arises and everything begins in that weariness tinged with amazement.25

For Gregor, however, this wake-up call comes too late. Along Camus’s line of reasoning, perhaps if Gregor could have snapped out of his trance of living like an insect, he would have transformed into a functional human. Catherine does not fare any better, considering she also wakes up but, like Gregor, does not know how to take advantage of the revelation. When Nelly and Catherine are talking the night Heathcliff leaves, Catherine says the obstacle to herself is in her forehead and heart, or “in whichever places the soul lives—in my soul, and in my heart, I’m convinced I’m
wrong!” But she clearly never wakes up, just like Gregor, considering she immediately catches “brain fever” afterward and later marries Edgar Linton, never having truly become authentic. She remains like Gregor’s picture: flat, predictable, and gilded.

Rationalization

Characters in *Wuthering Heights* and *The Metamorphosis*, entrapped by their individual struggles, attempt to rationalize their situations by unhelpful thought patterns and violence; however, because they are all victims of the absurd instead of being Camus’s fighting overcomers, none of the characters seem to have the ability of constructive rationalization. Matthew Kaiser argues of *Wuthering Heights* that “Brontë explores... a future-negating ethic of nonproductivity.” It should thus be no surprise that characters in *Wuthering Heights* never seem to get anywhere in their lives. The same can be said for Kafka’s *The Metamorphosis*.

For instance, Gregor’s physical transformation defies logic, yet Gregor is never shocked that he has mysteriously transformed into a bug. That could be attributed to the fact he has been a mere insect for years. In fact, his thoughts center around his family and their welfare despite his family caring little about him even when he was successful. “Thoughts like these, utterly futile in his current state, passed through his head.” Gregor is more concerned about his new appearance shocking the family than about his own state of mind. He may be physically locked within the body of an insect, yet his mind continues to function as a human’s. When his sister Grete assumes that removing his furniture would be beneficial for Gregor, as he would have more room to crawl around, Gregor realizes this act would be “at the price of simultaneously swiftly and completely forgetting his human past.” Gregor, as long as he lives, is forever caught. His metamorphosis is incomplete, as he cannot be fully insect or fully human.

In *Wuthering Heights*, characters have more insight into their states than Gregor has but fall into the same situation of attempting to rationalize their situations. Heathcliff knows the hurt that has been forced upon him by others but lacks a knowledge of how to escape it like Gregor. Heathcliff returns to *Wuthering Heights* to slowly torment the Lintons and Hindley Earnshaw for the cruel injuries they have dealt him. Unlike Gregor, who takes his injuries passively throughout the entire novella, Heathcliff “seemed a sullen, patient child; hardened, perhaps, to ill-treatment” and resorts to violence in his adulthood. Violence is a common theme that Brontë and Kafka both explore as part of rationalization.

Despite the overwhelming sense of violence, however, Graeme Tytler notes, “*Wuthering Heights* is conspicuous for violence of various kinds, it is useful to remember that... the most serious forms of violence... happen far less often in the narrative than some readers might be inclined to suppose or recall.” The violence
somehow does not seem so violent because of how absurd the situations are. Camus perhaps best states the sheer absurdity of the absurd: “Forever I shall be a stranger to myself…. Socrates’ ‘Know thyself’ has as much value as the ‘Be virtuous’ of our confessinals.” Camus’s point is there are no concrete truths to life, as everything is absurd, and thus, full comprehension of life and all it involves, including violence, seems somewhat laughable and illusionary.

Heathcliff, distraught over Catherine’s death, bashes his head against a tree, as he cannot live without his “life and soul.” Heathcliff, though, was by far not an innocent creature. At the time of Catherine’s death, he was gradually destroying Hindley for Hindley’s past cruelty and had already taken steps to destroy Edgar Linton by seducing the childish Isabella. Shortly after Catherine’s death, Hindley has a moment of clarity about Heathcliff’s damage and decides to kill Heathcliff with a pistol and knife; however, Heathcliff gains the upper hand and accidentally cuts one of Hindley’s arteries in the mêlée and then proceeds to “[kick] and [trample] on him, and [dash] his head repeatedly” before administering rudimentary first aid. Catherine has a history of self-violence, not so much on the physical level of self-harm but on the level of ostensibly contributing to her hysterical fits. She says at one point, “I’ll try to break their hearts by breaking my own.”

In The Metamorphosis, Gregor’s father, Mr. Samsa, is the most visibly cruel character, as he is described shoving Gregor back into his room, causing Gregor physical injury while brandishing a newspaper and stamping his foot. Later, he pelts Gregor with apples, leaving one lodged in Gregor’s exoskeleton. Grete is actually the cruelest character, however, considering that it is her care and attempted affection that initially keep Gregor in a happier state than he would have otherwise been under the circumstances. It is Grete who cleans his room and feeds him. Toward the end of the story, these tasks are left to a charwoman who likes to taunt Gregor. The family places unneeded belongings in Gregor’s room, demonstrating that they no longer want Gregor because he can no longer be useful to them. Grete finally repudiates Gregor by stating, “I am unwilling to utter my brother’s name before this creature, and therefore will say only: we have to try to get rid of it.” Grete’s rejection is what contributes to Gregor’s death that night, although physical causes are also to blame.

Catherine’s and Heathcliff’s deaths may also be looked at in a similar light. Catherine dies after childbirth, but her emotional state contributes to her death. She and Heathcliff meet for the last time in this world despite Edgar’s forbidding Heathcliff to return to Thrushcross Grange. After Catherine sinks down into an insensible state because of this meeting, Nelly states, “Far better that she should be dead, than a lingering burden, and a misery-maker to all about her.” Edgar mourns Catherine’s death, though, and raises Catherine’s child and namesake, Cathy. When Heathcliff dies, Hindley’s son Hareton mourns his death, although perhaps he should
not, as Heathcliff was a bad influence on his life, and Nelly states about Heathcliff’s grave, “I hope its tenant sleeps soundly.”

Camus states, “But practically I know men and recognize them by their behavior, but the totality of their deeds… by gathering together the sum of their consequences in the domain of intelligence, by seizing and noting all their aspects, by outlining their universe.”

All these characters—Catherine, Gregor, Grete, Heathcliff, and Hindley—are not limited to what has been thrust upon them. Throughout the course of both works, they choose their paths in life and then attempt to rationalize their states. The characters are known through their thoughts, behaviors, and actions. In truth, one cannot entirely blame Gregor’s father, Grete, Catherine, Hindley, or even Heathcliff, as they were trying to deal with impossible circumstances as well as they thought they could. Regardless of reasons for their behavior, no one picked efficient methods of rationalization. But honestly, do not we all do such at times?

**Authenticity vs. Inauthenticity**

Because of a lack of authenticity—the concept of whether a person’s decisions and actions are consistent with one’s personal identity and values or whether one is living according to social conformity—the only natural state for the worlds of *Wuthering Heights* and *The Metamorphosis* is the closed-in structure. Camus notes, “So long as the mind keeps silent in the motionless world of its hopes, everything is arranged and reflected in the unity of its nostalgia. But with its first move this world tumbles and cracks: an infinite number of shimmering fragments is offered to the understanding.” When the absurd is realized, the illusion of perfect life is shattered and then it is time to find out who cannot or does not know how to be authentic. Certainly, in the worlds of *Wuthering Heights* and *The Metamorphosis*, no one is authentic.

The best word to describe Catherine might be “disaster.” Her character was high-strung even as a young child, but the final break for her from finding any authenticity is when she is bitten by the Lintons’ dog. Catherine is “transformed into a very dignified person,” a doll whose ‘beautiful hair’ the Lintons ‘dried and combed,’ and whom they ‘wheeled’ dotingly, like a thing, to the fireplace…. They give her a husband. They give her a future. But she loses, in effect, her cosmos.” Even when she is older, her lack of “cosmos,” or essence, remains. Following the disappearance of Heathcliff, he vanishes for three years. The subsequent events after Heathcliff’s return reach a climax one evening when Edgar has enough of Heathcliff’s presence. He asks Catherine, “Will you give up Heathcliff hereafter, or will you give up me? It is impossible for you to be my friend, and his at the same time.” Catherine refuses to choose and finally be authentic for once, a decision that leads to her locking herself up in her room for three days because she cannot accept the absurd.
In the past, Catherine tried to be somewhat authentic by self-expression. She expresses herself uniquely on more than one occasion, including the exchange with Nelly in Volume I, Chapter IX but never goes further than mere expression of the self in words or rather, expression of the lack of her own self. Catherine tells Nelly:

[S]urely you and everybody have a notion that there is, or should be, an existence of yours beyond you. My great miseries in this world have been Heathcliff’s miseries, and I watched and felt each from the beginning; my great thought is living in himself. . . . Nelly, I am Heathcliff.  

Instead of attempting to find her own identity, Catherine expresses herself by identifying herself with another person. Understandably, Nelly is confused by this speech and annoyed, considering Catherine’s typically strange behavior and tempestuous ways. Catherine’s is a good example of the result of a life lived inauthentically.

Heathcliff, though, apparently can make his own decisions when he chooses to spend his life tormenting the Lintons and Hindley and cheating Hareton out of an education and inheritance while he, Heathcliff, is alive. Heathcliff’s son, Linton, is used as a mere tool to gain Thrushcross Grange from Cathy Linton since she is the heiress, but one must ask, even though Heathcliff is choosing his own essence in life, is it authentic? Heathcliff’s decisions never leave him any pleasure or peace of authenticity, so it can be assumed he is not living authentically. Neither do any of the other characters seem to be living authentically, although the novel closes with a bit of hope that Cathy and Hareton may be able to lead authentic lives.

Gregor’s inauthenticity comes from sacrificing himself for a family and company that use him only for his worth. His decisions eventually result in his death. Before he dies, though, Gregor attempts self-expression by walking on the walls and ceiling of his room. He explores the limitations and advantages of his insect body in ways he never explored his human body. The story ends with a foreboding that Gregor’s story may repeat itself through Grete. She had to sacrifice her dreams of attending a music conservatory because Gregor could no longer support them and may in the future sacrifice her essence for her parents or her future family when she marries.

An interesting thought pertaining to Brontë’s and Kafka’s existential theme of authenticity is that both authors are attempting self-expression by writing about characters with no authenticity. Emma Francis notes, “[I]t was precisely in terms of Brontë as a woman that I encountered problems in my analysis.” Of course, Francis’s main analysis was centered on Emily Brontë’s poetry, not her sole novel, but the same remains true. Kaiser referred to Brontë as “the great antiparent of nineteenth-century literature” because of Brontë’s rugged individualism when compared to other Victorian writers. Franz Kafka, although writing in the early twentieth century, also
utilized his own individualism in writing, although, as William Hubben points out, maybe Kafka’s existential symbolism should be taken with some reservations. Hubben notes that most of Kafka’s works were unfinished products at the author’s death and that “[p]erhaps it is this absence of solutions in his own life as well as in his work that makes these stories more symbolic than he may have intended.” The Metamorphosis is one of Kafka’s works that is thought to be in complete form, however, though even in complete form, it remains absent any true solution.

The point is that the authors of both works seem to have been attempting to locate their own sense of authenticity by writing about characters who completely lack authenticity. At any rate, regardless of the authors’ exact motives and personal struggles, readers will continue to search for some meaning in these works. After all, by no means is it certain that Brontë and Kafka intended for us to create all the myriad interpretations that have been used to explain Wuthering Heights and The Metamorphosis. Camus notes, “In this universe the work of art is then the sole chance of keeping his [humanity’s] consciousness and of fixing its adventures.”

Literary interpretations are like the absurd world: There is no external definability to human life; thus, anything produced by humans is by default essentially indefinable. Yet we must keep turning to the self-expressive art of art to continue to revolt amid the absurd.

Conclusion

Because life is absurd, according to Camus, he chose Sisyphus from Greek mythology to symbolize the plight of humans. Sisyphus notably was a tyrant who was punished by the gods in the afterlife to eternally roll a stone to the top of a hill, only to have the stone roll back down, for Sisyphus to repeat the cycle. Camus states, “The struggle itself towards the heights is enough to fill a man’s heart. We must imagine Sisyphus happy.” Sisyphus is still toiling against the absurdity of his situation; therefore, he is revolting against the absurd. Taking a lesson from Sisyphus, we must keep reading and analyzing to finally grasp the best interpretation of literature, to procure Camus’s “second reading.”

Through glimpses into Wuthering Heights and The Metamorphosis, it is possible to reflect upon Camus’s absurdity and the dangers of failing to recognize it. Characters are doomed to be enclosed in trapped worlds without any sense of authenticity. Characters are both confined and free simultaneously, as Gregor is confined to an insect body yet is free to attempt self-expression. His family is unsympathetic to his situation, especially by the end of the novella, something also seen in Wuthering Heights, although attitudes between characters tend to fluctuate more. The most striking part of both works is the imagery of locked doors and open windows. The open window provides a view into the world of the past and the unseen.
One can only hope that the characters beyond the events told in each story manage to unlock the doors and secure the windows, metaphorically speaking, considering these objects are associated with the downfall of the authentic self. After all, one must imagine Sisyphus happy.
Notes


4 Camus, 14.

5 For a few examples of this, see Brontë, 121, 162, 187, 335–36.

6 Emily Brontë, Wuthering Heights, ed. Pauline Nestor (London: Penguin Classics, 2008), 19. Catherine’s “large oak case” is also known as a “box-bed,” which was used in European homes, especially rural homes, from medieval times to the nineteenth century. The purpose of this style of bed was to create sleeping space in a house that otherwise did not have many bedrooms. This is apparently the case with Wuthering Heights, considering that Isabella is left with no proper place to sleep after marrying Heathcliff (143). For pictures of this style of bed, see https://en.wikipedia.org/wiki/Box-bed.

7 Brontë, 120.


10 Brontë, Wuthering Heights, 25.

11 Brontë, 126. “Brain fever,” as often seen in Victorian literature, as an illness following a severe emotional event, was an actual medical condition, although it was rooted in physical, not emotional, causes. It is thought to be either meningitis or encephalitis, neither of which was understood in the 1800s when Emily Brontë was writing Wuthering Heights. More than likely, the condition came to be used in literature after people in real life developed symptoms following emotional stress, and people connected the two because a medical explanation had not been discovered yet. (See https://daily.jstor.org/did-victorians-really-get-brain-fever/.)

12 Ibid.

13 Ibid.

14 Brontë, 336.

15 Brontë, 334.

16 Kafka, Metamorphosis, 32.

17 Kafka, 43–44.

18 Brontë, Wuthering Heights, 336.

19 Brontë, 288.

20 Vilma-Irén Mihály, “Hope and Hopelessness Through the Lens of Myths: A Comparison Based on

21 Camus, Myth of Sisyphus, 64.

22 Kafka, Metamorphosis, 4.

23 Brontë, Wuthering Heights, 53.

24 Kafka, Metamorphosis, 28.


28 Kafka, Metamorphosis, 23.

29 Kafka, 27.

30 Brontë, Wuthering Heights, 38.


32 Camus, Myth of Sisyphus, 19.

33 Brontë, Wuthering Heights, 169.

34 Brontë, 178–79.

35 Brontë, 116.

36 Kafka, Metamorphosis, 16.

37 Kafka, 32.

38 Kafka, 41.

39 Brontë, Wuthering Heights, 164.

40 Brontë, 336.

41 Camus, Myth of Sisyphus, 11.

42 Camus, 18.


44 Brontë, Wuthering Heights, 117.

45 Brontë, 82.


50 Camus, Myth of Sisyphus, 94.
Camus, 123.
THE INFLUENCE OF SPECIFIC DIAGNOSIS AND TRAIT SELF-RELEVANCE ON POSITIVE AND NEGATIVE PERCEPTIONS OF INDIVIDUALS WITH MENTAL ILLNESS

SABRINA L. STOCK, BUTLER UNIVERSITY
AEDAN S. LEVY, BUTLER UNIVERSITY
MENTOR: TARA LINEWEAVER

Abstract

We investigated how participants perceived an individual with either depression, schizophrenia, or no disorder on negative and positive traits and then examined how these perceptions linked to desire for social distance (DSD), trait self-relevance, positive former contact with a person with mental health problems, and continuum beliefs. One hundred fifty undergraduates completed an online survey that described a vignette character with depression (VCWD), vignette character with schizophrenia (VCWS), or control character with no mental disorder (CC) and assessed participants’ perceptions of the character’s traits, DSD from the vignette character, perceptions of one’s own traits, and demographic characteristics. In general, the VCWD was viewed as thoughtful, unpredictable, and apathetic. The VCWS was viewed as interesting but less kind as well as more dangerous, irresponsible, and unpredictable than other characters. DSD correlated with perceived dangerousness, irresponsibility, and unpredictability. Because negative traits are strong predictors of DSD, participants showed more DSD from the VCWS than other characters. Neither positive former contact nor continuum beliefs correlated with DSD. Finally, high trait self-relevance increased the likelihood of perceiving one’s own traits in a VCWD but not a VCWS, suggesting participants saw less of themselves in someone with schizophrenia than someone with depression. Together, these results suggest that positive and negative perceptions of depression and schizophrenia exist but people identify less with and are less willing to socialize with someone diagnosed with schizophrenia than with depression.

Perceptions of mental illnesses are critical to addressing mental health. Stigma, a useful measure of negative perceptions, is best defined as negative attitudes that are associated with labeling, stereotyping, separating, and discriminating (Link & Phelan, 2001) and can be measured through desire for social distance as a behavioral component (Buckwitz et al., 2021). Trivialization is best defined as behaviors that portray illnesses as easy to manage, treat, or live with (Robinson et al., 2019). One
study, which compared stigma against mental versus physical illnesses by analyzing thousands of tweets on the topics from the social media platform Twitter, found that mental health conditions are, on the whole, stigmatized and trivialized significantly more than are physical conditions (Robinson et al., 2019). The most stigmatized of the ten analyzed illnesses was schizophrenia. Notably, schizophrenia was stigmatized more than it was trivialized, while the opposite held true for depression. Additionally, of the mental illnesses, depression was the target of the least stigma and the second least trivialization. Another study, which analyzed the contents of tweets with the hashtags #depression or #schizophrenia, found that people are more supportive of depression than of schizophrenia and that stigma against schizophrenia is far greater comparatively (Reavley & Pilkington, 2014).

Past research has established that understanding social perceptions of individuals with mental illnesses is important because three of the most prominent factors that predict social rejection of people with mental disorders are perceived personal responsibility, dangerousness, and rarity of mental illness (Anderson et al., 2015; Feldman & Crandall, 2007). When examined in the context of depression and schizophrenia, people tend to believe that individuals with schizophrenia are dangerous, whereas they perceive depression to be a result of personal weakness (Robinson et al., 2019), traits which increase desire for social distance (Anderson et al., 2015; Feldman & Crandall, 2007; Yap et al., 2014). Additionally, perceived unpredictability has also been associated with greater desire for social distance against people with a mental illness (Yap et al., 2014).

At the same time, several factors seem to mitigate negative perceptions of mental illnesses and thus reduce stigma and desire for social distance. First, gender seems to play an important role. Compared to men, women view depression as significantly more common, more unavoidable, and less the person’s fault (Anderson et al., 2015); thus, women seem to have fewer and weaker stigmatizing attitudes compared to men. Second, continuum beliefs versus dichotomous beliefs around mental illness play a role in desire for social distance and stigma as well (Buckwitz et al., 2021). As described by Buckwitz et al., a dichotomous perspective posits that having symptoms of mental illness places an individual into a separate, distinct category: either one is mentally ill or one is “normal.” A continuum perspective, in contrast, posits that everyone can experience symptoms of mental illness from time to time. For example, under this perspective, an individual who is showing periods of low mood (a symptom of depression) does not necessarily have depression. Instead, their current situation, environment, or state of being may simply amplify the feeling of low mood to a more pronounced but temporary state. Continuum beliefs are associated with less desire for social distance, less stigma, and less belief of dangerousness (Buckwitz et al., 2021; Peter et al., 2021). Finally, simply interacting with a person with a mental illness in a positive way can influence perceptions of mental illnesses. Having positive former contact with someone who has a mental illness is associated with
increased continuum beliefs, less stigma, and less desire for social distance (Buckwitz et al., 2021).

All of these associations may be important in the context of a pandemic-stricken world. Amidst the COVID-19 pandemic, symptoms of mental illness have skyrocketed: 40.9% of people reported an adverse mental or behavioral condition, with 30.9% reporting anxiety or depression symptoms, 26.3% reporting trauma- and stressor-related disorder (TSRD) symptoms related to COVID, and 10.7% having seriously considered suicide (Czeisler et al., 2020). Young adults have undoubtedly suffered a disproportionately greater negative impact during the pandemic, as 74.9% of 18–24-year-olds reported an unhealthy mental or behavior condition (Czeisler et al., 2020). Given the sheer weight of these statistics, it is likely that most people know at least one loved one (friend, family, or coworker, for instance) presenting symptoms of mental illness. Because positive former contact with someone experiencing a mental illness is associated with decreased desire for social distance (Buckwitz et al., 2021), people may desire less social distance from people with mental illnesses overall because of their interactions with struggling loved ones.

Collectively, this increase in mental illness symptoms and presence on social media may have led to decreased negative perceptions and possibly the emergence of positive ones. As more people experience symptoms of mental illness or are in contact with people experiencing symptoms, these mental illnesses may be perceived to decrease in rarity, a factor that has been shown to increase desire for social distance (Feldman & Crandall, 2007). Furthermore, although no literature exists on the topic, continuum beliefs during the pandemic may have increased, given that so many people without formal mental illness diagnosis displayed symptoms of one. Finally, because 90% of individuals with self-identified mental health issues want to participate in online social media programs designed to help people cope with mental health symptoms (Naslund et al., 2019), people struggling during the pandemic with their own mental health symptoms may have turned to social media to receive destigmatizing information and support. Although a wealth of literature on the negative perceptions of mental illnesses exists, little research focuses on the ways in which mental illnesses are perceived to be positive or to contribute to a person’s character in a beneficial way, and hardly any literature exists on either negative or positive perceptions postpandemic. One of the most notable, if not the only notable, positive connotations with mental illness is increased creativity, which has long been contested as an associated characteristic of mental illnesses (Kaufman et al., 2006). This stereotype of the mad genius may bring to mind famous and respected artists such as Van Gogh, Cobain, or Hemingway, all mentally ill creatives. Individuals are more likely to endorse the mad genius stereotype (defined in the study’s questionnaire with the terms mental illness, schizophrenia, depression, and bipolar) if they consider themselves creative, regardless of whether they actually are creative on objective measures (Kaufman et al., 2006). In other words, this result suggests that individuals
who consider creativity a personality trait relevant to themselves are more likely to believe that mental illness is associated with creativity. The current study expanded on the past literature by focusing on how individuals may associate both positive and negative traits with mental illnesses. This study additionally examined how these associations relate to desire for social distance, trait self-relevancy, continuum beliefs, and positive former contact with a person with mental health problems. We hypothesized four relationships: First, participants would perceive a vignette character with depression more positively than a vignette character with schizophrenia. Second, participants would desire the greatest social distance from a vignette character with schizophrenia. Third, desire for social distance from a vignette character with a mental illness would be associated with continuum beliefs, positive former contact with a person with a mental health problem, and perceived levels of dangerousness, irresponsibility, and unpredictability. Fourth, participants who view themselves as more creative would view a vignette character with a mental illness as more creative.

Method

Participants

Participants for this study were recruited via Sona (an online site through which students can participate in studies for extra credit in classes), the Butler Honors program listserv (a list of emails of all Butler students currently enrolled in the Butler Honors program), and personal connections. One hundred fifty undergraduate students took part in this study. Demographic characteristics of the participants are summarized in Table 1. Participants in the three conditions did not differ from each other significantly in their age ($F(2, 147) = .390, p = .678, \eta_p^2 = .005$), continuum beliefs ($F(2, 147) = .21, p = .81, \eta_p^2 = .003$), or number of psychology courses that they had taken at the high school or college level, $F(2, 147) = 0.97, p = .381, \eta_p^2 = .01$. Additionally, the distribution of gender ($\chi^2(n = 150, df = 8) = 14.70, p = .065$), race ($\chi^2(n = 150, df = 10) = 13.60, p = .192$), past diagnosis with a mental illness ($\chi^2(n = 150, df = 2) = 3.90, p = .142$), participation in therapy ($\chi^2(n = 150, df = 2) = 0.65, p = .723$), and relationship with someone in treatment for a mental health problem ($\chi^2(n = 150, df = 4) = 3.93, p = .416$) was statistically equivalent across groups.
Table 1. Demographic Characteristics of Participants Who Rated the Vignette and Control Characters

<table>
<thead>
<tr>
<th>Demographic trait</th>
<th>Vignette character with depression (n = 52)</th>
<th>Vignette character with schizophrenia (n = 50)</th>
<th>Control character (n = 48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (SD)</td>
<td>19.92 (2.52)</td>
<td>20.16 (4.027)</td>
<td>19.71 (1.11)</td>
</tr>
<tr>
<td>Gender</td>
<td>88.0% Female</td>
<td>82% Female</td>
<td>79% Female</td>
</tr>
<tr>
<td></td>
<td>7.7% Male</td>
<td>8% Male</td>
<td>18.8% Male</td>
</tr>
<tr>
<td></td>
<td>0% Nonbinary</td>
<td>8% Nonbinary</td>
<td>2% Nonbinary</td>
</tr>
<tr>
<td></td>
<td>0% Other identity</td>
<td>2% Other identity</td>
<td>0% Other identity</td>
</tr>
<tr>
<td></td>
<td>3.8% Prefer not to answer</td>
<td>0% Prefer not to answer</td>
<td>0% Prefer not to answer</td>
</tr>
<tr>
<td>Race</td>
<td>90.1% White</td>
<td>88% White</td>
<td>81.3% White</td>
</tr>
<tr>
<td></td>
<td>0% Black</td>
<td>0% Black</td>
<td>4.2% Black</td>
</tr>
<tr>
<td></td>
<td>5.8% Asian</td>
<td>4% Asian</td>
<td>2.1% Asian</td>
</tr>
<tr>
<td></td>
<td>1.9% Hispanic</td>
<td>0% Hispanic</td>
<td>8.3% Hispanic</td>
</tr>
<tr>
<td></td>
<td>1.9% Multiracial</td>
<td>6% Multiracial</td>
<td>2.1% Multiracial</td>
</tr>
<tr>
<td></td>
<td>0% Other</td>
<td>2% Other</td>
<td>0% Other</td>
</tr>
<tr>
<td>Mental illness diagnosis</td>
<td>61.6% Yes</td>
<td>42% Yes</td>
<td>52.1% Yes</td>
</tr>
<tr>
<td>Mean continuum beliefs (SD)</td>
<td>38.4% No</td>
<td>58% No</td>
<td>47.9% No</td>
</tr>
<tr>
<td>Mean number of psychology courses (SD)</td>
<td>16.71 (2.18)</td>
<td>16.44 (2.16)</td>
<td>16.56 (1.97)</td>
</tr>
<tr>
<td></td>
<td>3.56 (1.49)</td>
<td>3.54 (1.47)</td>
<td>3.90 (1.29)</td>
</tr>
</tbody>
</table>

Materials

Vignettes

The vignettes, which were modeled after vignettes used in a study of ingroup and outgroup perceptions of people with mental illnesses (Douglas & Sutton, 2011), consisted of a description of a 21-year-old college student. All vignettes provided a description of a person who participates in the common activities of that age demographic, such as going to the movies, hanging out with friends, and going to the recreation center. The gender of the vignette character was purposefully left ambiguous by using a gender-neutral name (S. Hedron) and they/them pronouns. The
control vignette included a statement that the character has never had any physical or mental disorders and is generally healthy. The target vignettes included a statement that the character has had either schizophrenia or depression for the last three years. (See appendix.) The individuals described in these vignettes are referred to as the control character (CC), the vignette character with schizophrenia (VCWS), and the vignette character with depression (VCWD).

**Stock Characteristics of Other Scale**

Developed for the present study, this scale measured the degree to which the participant perceived the vignette character as having each of nine characteristics (five positive, four negative) relative to the average person using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). An example item includes “Compared to the average person, S. Hedron is more kind.” The five positive traits included creativity, intelligence, kindness, thoughtfulness, and interestingness, and the four negative traits included dangerousness, irresponsibility, unpredictability, and apathy.

**Social Distance Scale (Bogardus, 1925)**

This 7-item self-report questionnaire, typically used to identify social rejection against specific groups, measured the degree to which the participant would desire to avoid or socially distance from an individual on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Questions focused on various social settings and relationships, including friendship, romance, and workplace. Example items include “I would like this person to come and work at the same place I do” and “I would like this person to be a close personal friend.” Responses were reverse-scored so higher scores indicated a greater desire for social distance.

**Stock Characteristics of Self Scale**

Developed for the present study, this scale measured the degree to which the participant perceived themselves as having nine characteristics relative to the average person, using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). The characteristics mirrored those from the Stock Characteristics of Other Scale. An example item includes “Compared to the average person, I am more kind.”

**Beliefs About Illness Scale (Norman et al., 2008)**

This 4-item self-report questionnaire measured the degree to which the participant believed mental illnesses lay on a continuum, using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). This questionnaire included three items from Norman et al. (2008) and one reverse-scored item created by the author. An
example item includes “Normal people can have some of the symptoms of _____.” Higher scores indicated a greater belief that mental illnesses lie along a continuum rather than being a dichotomous category.

Demographic Questionnaire

Participants provided information regarding their general demographics, such as race, gender, age, and class standing. Further demographic questions asked whether the participant had ever been diagnosed with a mental illness, participated in therapy, known someone who had sought treatment for a mental health condition (and, if so, if their interactions with this person had been primarily positive or negative), or participated in psychology courses (and, if so, how many). The questions measuring positive former contact were replicated from Buckwitz et al. (2021).

Procedure

Participants completed the anonymous online survey via the platform Qualtrics. The survey began with an informed consent, and then participants read the vignette of the fictional character. They completed the Stock Characteristics of Other Scale, the Social Distance Scale, the Stock Characteristics of Self Scale, and the Beliefs About Illness Scale. Finally, they completed the demographic questionnaire. After they finished the survey, participants were thanked and received extra credit in their psychology courses, if applicable.

Results

Perceived Positive and Negative Traits of Vignette Characters by Condition

To evaluate the perceived positive and negative traits of vignette characters in each condition, we utilized two multivariate analyses of variances (MANOVAs), one evaluating positive traits and one negative. Condition, the specific vignette that the participant read, was the between-subjects independent variable.

The condition main effect reached significance for positive traits, $F(10, 286) = 7.74, p < .001, \eta^2_p = .21$ (Figure 1). An examination of the univariate analyses indicated a significant effect of condition for the traits of interesting ($F(2, 147) = 12.98, p < .001, \eta^2_p = .15$), thoughtful ($F(2, 147) = 4.51, p = .013, \eta^2_p = .06$), and kind ($F(2, 147) = 9.17, p < .001, \eta^2_p = .11$) but not creative ($F(2, 147) = 2.61, p = .077, \eta^2_p = .03$) or intelligent, $F(2, 147) = 2.12, p = .124, \eta^2_p = .03$. Follow-up analyses with Tukey’s test indicated that the VCWD was viewed as more thoughtful than the VCWS ($p = .004$) and CC ($p = .046$).
Compared to the VCWD and CC, the VCWS was viewed as more interesting but less kind (all $ps \leq .001$).

The condition main effect also reached significance for negative traits, $F(8, 288) = 19.42, p < .001, \eta_p^2 = .35$ (Figure 2). Univariate analyses showed a significant difference based on condition for all four negative traits: dangerous ($F(2, 147) = 24.36, p < .001, \eta_p^2 = .25$), unpredictable ($F(2, 147) = 57.68, p < .001, \eta_p^2 = .44$), apathetic ($F(2, 147) = 11.66, p < .001, \eta_p^2 = .14$), and irresponsible, $F(2, 147) = 12.49, p < .001, \eta_p^2 = .15$. Follow-up analyses with Tukey’s test indicated that the VCWS was viewed as more dangerous, irresponsible, and unpredictable than both other vignette characters (all $ps < .001$). The VCWD was viewed as more apathetic than the VCWS ($p = .003$) and CC ($p < .001$), as well as more unpredictable than the CC ($p < .001$).
Perceived Desire for Social Distance from Vignette Characters by Condition

To examine whether perceived desire for social distance (DSD) depended on vignette character descriptions, we utilized a univariate analysis of variance with condition as an independent variable and DSD as a dependent variable. DSD differed significantly across conditions, $F(2, 147) = 4.51, p = .013, \eta_p^2 = .06$ (Figure 3). Follow-up analyses with Tukey’s test revealed that participants desired more social distance from the VCWS than the VCWD ($p = .007$) or the CC ($p = .016$) but that DSD did not differ between the VCWD and CC ($p = .810$).
Predictors of or Factors Related to DSD

We ran correlations to evaluate the relationship between the ratings of the vignette characters on the five positive and four negative traits and desire for social distance. Participants had greater DSD when vignette characters were perceived as more dangerous \((r = .347, p < .001)\), irresponsible \((r = .405, p < .001)\), and unpredictable \((r = .209, p = .010)\) than the average person. Additionally, perceived thoughtfulness, kindness, and intelligence of the vignette characters reduced DSD. See Table 2 for complete list of traits correlated with DSD.

Table 2. Relationships Between Perceived Positive (left column) and Negative (right column) Traits of the Vignette Character and DSD

<table>
<thead>
<tr>
<th>Trait</th>
<th>Positive traits</th>
<th>Negative traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interestingness</td>
<td>–.114</td>
<td>Dangerousness</td>
</tr>
<tr>
<td>Thoughtfulness</td>
<td>–.219**</td>
<td>Unpredictability</td>
</tr>
<tr>
<td>Creativity</td>
<td>–.069</td>
<td>Apathy</td>
</tr>
<tr>
<td>Kindness</td>
<td>–.228**</td>
<td>Irresponsibility</td>
</tr>
<tr>
<td>Intelligence</td>
<td>–.183*</td>
<td></td>
</tr>
</tbody>
</table>

*\(p < .05\).  **\(p < .001\).
We ran correlations to evaluate the relationship between continuum beliefs and DSD in each and across conditions. The participants’ continuum beliefs did not significantly correlate with DSD in the control condition ($r = -0.040, p = 0.788$), in schizophrenia condition ($r = -0.176, p = 0.221$), in the depression condition ($r = -0.107, p = 0.449$), or across all conditions ($r = -0.126, p = 0.126$).

To examine whether a participant’s positive former contact with a person with mental health problems was associated with DSD from vignette characters in each condition, we first divided the sample into participants who knew someone who had been in treatment for a mental health problem and those who did not. We then identified participants who had had a positive experience with a person who had been in treatment. Because only 8 participants across all conditions had negative former contact, we did not include these participants in any analyses.

Once the sample had been divided between those who had positive former contact ($n = 105$) and those who had no former contact ($n = 30$), we utilized a univariate analysis of variance with condition and former contact as independent variables and DSD as the dependent variable. The main effect of former contact did not reach significance, $F(2, 134) = 0.259, p = 0.772, \eta^2_p = 0.004$.

**Perceived Trait Self-Relevance of Vignette Characters by Condition**

To examine the association between trait self-relevance and trait perceptions of the vignette characters, we ran correlations for each trait and by condition. Participants in the depression condition who viewed themselves as more interesting were more likely to view the VCWD as more interesting ($r = -0.458, p < 0.001$). This pattern also emerged for dangerousness ($r = 0.424, p = 0.002$), unpredictability ($r = 0.300, p = 0.031$), and apathy ($r = 0.456, p < 0.001$). No self-relevance correlations reached significance for participants in the schizophrenia condition, but control condition participants’ self-rated intelligence ($r = 0.374, p = 0.009$), dangerousness ($r = 0.313, p = 0.030$), and apathy ($r = 0.287, p = 0.048$) correlated with their ratings of the CC on the corresponding traits. See Table 3.
Table 3. Relationships Between Ratings of Self-Relevant Traits and Ratings of Vignette Character Traits by Condition

<table>
<thead>
<tr>
<th>Character trait</th>
<th>Self-rating of same trait</th>
<th>Vignette character with depression</th>
<th>Vignette character with schizophrenia</th>
<th>Control character</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive traits</td>
<td>Negative Traits</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interestingness</td>
<td>.458**</td>
<td>.213</td>
<td>.013</td>
<td></td>
</tr>
<tr>
<td>Thoughtfulness</td>
<td>.137</td>
<td>.167</td>
<td>.154</td>
<td></td>
</tr>
<tr>
<td>Creativity</td>
<td>.153</td>
<td>−.120</td>
<td>.052</td>
<td></td>
</tr>
<tr>
<td>Kindness</td>
<td>.141</td>
<td>.224</td>
<td>.091</td>
<td></td>
</tr>
<tr>
<td>Intelligence</td>
<td>.222</td>
<td>−.086</td>
<td>.374**</td>
<td></td>
</tr>
<tr>
<td>Dangerousness</td>
<td>.424**</td>
<td>.006</td>
<td>.313*</td>
<td></td>
</tr>
<tr>
<td>Unpredictability</td>
<td>.300*</td>
<td>−.014</td>
<td>.075</td>
<td></td>
</tr>
<tr>
<td>Apathy</td>
<td>.456**</td>
<td>−.134</td>
<td>.287*</td>
<td></td>
</tr>
<tr>
<td>Irresponsibility</td>
<td>.074</td>
<td>.058</td>
<td>−.019</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.  **p < .01.

Role of Trait Self-Relevance in Perceptions of Vignette Characters

To determine whether trait self-relevance influenced ratings of the vignette characters on positive and negative traits, we wanted to ensure that trait self-perceptions were evenly distributed across the three vignette character conditions. Self perceptions of eight of the nine traits were equally balanced across conditions, all Fs (2, 147) ≤ 1.54, all ps ≥ .218. The only trait that varied across groups was self-perceived apathy, $F(2, 147) = 3.20, p = .044$. Participants in the VCWD condition rated themselves as significantly more apathetic ($M = 3.02, SD = 1.00$) than those in either the VCWS condition ($M = 2.48, SD = 1.18$) or those in the control condition ($M = 2.67, SD = 1.58$).

To determine whether this difference influenced the way participants rated vignette characters, we ran a univariate analysis of variance with condition as the independent variable, apathy ratings of S. Hedron as dependent variable, and apathy self-ratings as covariate. Self-rated apathy was significantly related to ratings of apathy in S. Hedron $F(1, 146) = 6.60, p = .011, \eta_p^2 = .043$; however, the effect of condition remained significant ($F(2, 146) = 10.06, p > .001, \eta_p^2 = .121$) even with the covariate included in the analysis.
Discussion

The goal of this study was to assess individuals’ positive and negative perceptions of depression and schizophrenia and how desire for social distance relates to these perceptions, trait self-relevancy, continuum beliefs, and prior former contact with a person with mental health problems. This study had four main hypotheses. First, we expected that participants would perceive the VCWD more positively than the VCWS, who would be seen most negatively of all the vignette characters. Second, we anticipated that, if the first hypothesis was supported, participants would have a greater desire for social distance from the VCWS than the other characters. Third, we hypothesized that continuum beliefs, positive former contact with a person with mental health problems, and perceived levels of dangerousness, unpredictability, and irresponsibility would be associated with desire for social distance from a vignette character with a mental illness (i.e., the VCWD or VCWS). Fourth, we expected that participants who viewed themselves as more creative would view the VCWD and VCWS as more creative than the CC.

Our results support our first hypothesis. Both the VCWS and VCWD were associated with both positive and negative traits, but the VCWD and CC were seen more positively than was the VCWS. Specifically, the VCWD was perceived as the most thoughtful but also the most apathetic of the vignette characters, whereas the VCWS was perceived to be the most interesting but the most dangerous, unpredictable, and irresponsible of the vignette characters. Such results support past literature, which has found that depression has one of the lowest rates of stigma associated with it whereas schizophrenia has one of the highest (Robinson et al., 2019). Additionally, neither the VCWD nor VCWS was perceived as being any more creative or intelligent than the control, contradicting the “mad genius” stereotype of mental illness (Kaufman et al., 2006). Although these results suggest that people with mental illnesses are viewed, in general, more negatively than people without, they also suggest that perceptions of mental illnesses include positive aspects.

Our second hypothesis was also supported. The participants predictably desired the most social distance from the VCWS. This result is consistent with past literature that found schizophrenia is greatly associated with dangerousness, irresponsibility, and unpredictability, which are traits that have predicted DSD in past studies (Feldman & Crandall, 2007; Yap et al., 2014). Because DSD can be understood as a behavioral component of stigma (Buckwitz et al., 2021), these results further suggest that participants had the greatest stigma against people with schizophrenia; however, participants did not desire significantly more social distance from the VCWD than the CC. This result is unexpected, given that people tend to perceive depression as a result of personal weakness (Robinson et al., 2019; Reavley & Pilkington, 2014), which is a predictor of DSD (Feldman & Crandall, 2007). This may suggest that less stigma is associated with depression postpandemic.
Our third hypothesis—which predicted that specific negative traits (dangerousness, unpredictability, and irresponsibility), continuum beliefs, and positive former contact would be associated with DSD—was partially supported. Results suggest that dangerousness, unpredictability, and irresponsibility did significantly predict DSD; however, continuum beliefs did not correlate with desire for social distance, which refutes part of our third hypothesis and, in fact, runs contrary to prior research (Buckwitz et al., 2021; Peter et al., 2021). Similarly, we found that positive former contact did not ameliorate DSD, which also differs from prior literature (Buckwitz et al., 2021). Because this study replicated Buckwitz et al.’s measurement and analyses of positive former contact, this deviation from past literature cannot be explained by a change in methodology. A possible explanation for why neither continuum beliefs nor positive former contact related to DSD is that, in the current study, the vignette characters were described as college students, which was not the case with the vignettes of either Douglas & Sutton (2011), whose work these vignettes were modeled after, or Buckwitz et al. (2021). As such, the participants, who were also college students, may have perceived greater similarity between themselves and the characters, resulting in a lower DSD. Our fourth hypothesis was not supported; creativity trait self-relevance did not relate to creativity trait perception of the vignette characters with mental illness, contradicting prior research finding that the more creative people viewed themselves, the more likely they were to believe individuals with mental illnesses were also creative (Kaufman et al., 2006); however, for other traits, trait self-relevance did increase the likelihood of perceiving one’s own characteristics in the CC or the VCWD but not in the VCWS. Specifically, people who saw themselves as more dangerous or apathetic were more likely to rate the VCWD or CC as more dangerous or apathetic. Individuals who saw themselves as more interesting saw the VCWD as more interesting, and individuals who saw themselves as more intelligent saw the CC as more intelligent. None of the correlations reached significance for the VCWS. These results suggest that participants saw less of themselves in someone with schizophrenia than in someone with depression or no diagnosis.

An unexpected finding not related to our hypotheses was that participants who read about the VCWD were higher in self-rated apathy than were participants in the other two conditions. Our data do not allow us to determine the direction of this relationship. It is possible that reading about the VCWD encouraged participants to rate their own apathy higher because they may have felt more comfortable admitting to their own apathy in the context of having read about a character who could also be perceived as apathetic. On the other hand, the finding that those who perceived themselves as more apathetic also rated S. Hedron as more apathetic could be an effect of trait self-relevancy similar to that documented in Kaufman et al. (2006) for creativity. It is also interesting to note that participants’ self-ratings on the other traits did not differ depending on the condition that they were assigned to. As an example,
participants who read about the VCWS did not rate themselves higher in dangerousness or unpredictability. Additionally, trait self-ratings of apathy did not fully explain the differences in how S. Hedron was perceived on that same characteristic based on whether the vignette character was described as being diagnosed with depression, schizophrenia, or no disorder.

This study contributes to the current literature in a number of ways. By better understanding the different traits associated with mental illnesses, organizations concerned with mental health and illness can better engage with biases that people have against, and foster a more welcoming environment for, individuals with mental illnesses. Our results suggest that people see themselves in the VCWD or CC but not in the VCWS; thus, people may see the VCWS as different and less like them. Knowing the ways in which people relate or do not relate to those with specific mental illnesses is crucial toward understanding how to manage and reduce stigma. Educational efforts regarding mental illnesses, particularly schizophrenia, may benefit from highlighting the similarities between those with a mental illness and those without.

Additionally, these results expand on the research of Kaufman et al. (2006) by identifying dangerousness, apathy, and interestingness as possible other self-relevancy traits that relate to perceptions of people with mental illnesses. Furthermore, an individual’s perception of their own intelligence played a role in their likelihood of perceiving a healthy person as intelligent, which suggests that self-relevancy effects are not limited to perceptions of people with mental illness; however, both the participants and the vignette character were college students, which may have increased self- and vignette ratings of intelligence. Together, these results suggest that a person’s positive and negative perceptions of other individuals, including those with depression but not those with schizophrenia, are influenced by a person’s perception of themselves.

Limitations and Future Directions

Some limitations of this study do exist. First, because participants were primarily female, we cannot determine if these results can be generalized across genders. According to Anderson et al. (2015), women are more likely to view depression as less rare and less the fault of the person with depression, which are beliefs that decrease DSD (Feldman & Crandall, 2007). As such, perceptions may have been more positive overall and desire for social distance lower in this study compared to past studies because it included fewer male participants. Second, 91.3% of participants across conditions had taken one or more psychology courses in the past ($M = 3.66, SD = 1.423$). Although the number of psychology courses was balanced across conditions, it is possible that, because of their background in the topic, some participants had a greater understanding of or empathy for people with mental illnesses, resulting in more positive perceptions overall. Third, given that the vignettes
explicitly stated that the character had depression or schizophrenia, people may have been conscious of the nature of this study and may have altered their answers in order to conform to social desirability. For example, they may have hesitated to admit negative views of a person with a mental illness, especially considering that many participants have taken psychology courses and thus may be engaged in the psychology department. Furthermore, the vignettes associated S. Hedron with a specific diagnosis (i.e., depression or schizophrenia) instead of describing symptoms associated with these disorders; thus, our outcomes reflect stigma associated only with the name depression or schizophrenia. We constructed the vignettes this way to ensure that participants provided their perceptions of the specific disorders that we were investigating rather than incorrectly reported perceptions for a different disorder with overlapping symptoms. Additionally, describing symptoms of diagnoses in the vignettes may have explicitly included traits, such as a person with depression being apathetic, which we wanted to investigate as outcomes. Fourth, all nine traits were presented in the same order across questionnaires, which may have led to order or carryover effects, which we did not analyze in this study. Finally, to limit the number of calculations we ran, we did not investigate how continuum beliefs were associated with perceived dangerousness or positive former contact, relationships which have been found to be significant in prior studies (Buckwitz et al., 2021).

Future studies should investigate different negative and positive traits in relation to perceptions of those with a mental illness to further distinguish how these perceptions increase or decrease DSD and stigma in general. Additionally, having more diverse participants may help determine how a person’s social identities, such as race, gender, and religion, influence their perceptions of people with mental illnesses. Furthermore, more investigation of the role of trait self-relevancy in forming perceptions of those with diagnosed mental illness may help psychologists better understand who is prone to form positive versus negative perceptions of those with mental illness. Finally, including vignette characters who are more varied in social identities, such as age, occupational status (e.g., college student vs. working a job), race, or gender, could help determine how these individual differences affect people’s perceptions of individuals with mental illness.

Conclusion

Modern-day perceptions of mental illnesses are not strictly negative. This study suggests that mental illnesses, specifically depression and schizophrenia, are associated not only with negative traits but also with positive characteristics. People generally perceive more negative traits in someone with schizophrenia, resulting in greater DSD, and hold more balanced views of someone with depression, who is typically seen as more similar to a healthy individual.
Furthermore, trait self-relevancy plays a role in a person’s perception of an individual with depression but not in their perception of an individual with schizophrenia, which may have contributed to the VCWD being perceived more positively than the VCWS. By better understanding the factors that influence perceptions of and desire for social distance from individuals diagnosed with mental illnesses, mental health organizations can more effectively address negative perceptions that people hold against individuals with mental illnesses.
References


Intervention in Psychiatry, 13(3), 405–413. 
http://dx.doi.org/10.1111/eip.12496


http://dx.doi.org/10.1017/S0033291721000854

http://dx.doi.org/10.7717/peerj.647

http://dx.doi.org/10.1007/s00127-018-1571-5

Appendix

Vignettes

Control Character

S. Hedron is a 21-year-old student. Hedron enjoys going to the movies, hanging out with friends, and going to the recreation center in their spare time. They have never had any physical or mental disorders and are generally healthy.

Vignette Character with Schizophrenia

S. Hedron is a 21-year-old student. Hedron enjoys going to the movies, hanging out with friends, and going to the recreation center. They have never had a physical disorder but have had schizophrenia for the last three years.

Vignette Character with Depression

S. Hedron is a 21-year-old student. Hedron enjoys going to the movies, hanging out with friends, and going to the recreation center. They have never had a physical disorder but have had depression for the last three years.
COMPARING THE VISUOSPATIAL ABILITIES OF ALZHEIMER’S PATIENTS AND HUNTINGTON’S PATIENTS

ALYSA ARALIS, BUTLER UNIVERSITY
EMMA BAUMAN, BUTLER UNIVERSITY
MENTOR: TARA LINEWEAVER

Abstract

Visuospatial deficits emerge from the pathophysiology of both Huntington’s disease (HD) and Alzheimer’s disease (AD), but few studies have directly compared the types of visuospatial deficits associated with these disorders. We directly compared the mental rotation abilities of HD ($n = 18$) and AD ($n = 18$) patients relative to age-matched control groups (young healthy controls: $n = 20$; older healthy controls: $n = 20$). Participants completed the nonrotational and rotational components of the Right-Left Orientation Test (RLOT), Luria, Money Road Map Test (MRMT), and Stick Construction Test (SCT). Participants also indicated the strategy they used to complete each: personal rotation (mentally rotating themselves through space to take a new perspective), extrapersonal rotation (mentally rotating test materials through space), or a nonrotational strategy (e.g., using knowledge about spatial relationships). Results showed that HD and AD patients were equally impaired relative to their control groups on rotational, but not nonrotational, components of all four tests. The four groups did not differ significantly in their self-reported use of strategies on the RLOT or Luria, but group differences emerged on the MRMT and SCT. Controls were more likely to report using personal rotation than either extrapersonal or other strategies on both MRMT and SCT, but their strategy choice did not affect their performance on either test. For patients, HD and AD groups reported using personal strategies as commonly as their respective healthy controls during SCT, but HD patients were less likely than their control group to use personal rotation on the MRMT. Utilization of a personal rotation strategy by patients resulted in better performance on the MRMT, but not on the SCT. Together, results suggest that many nonrotational visuospatial skills are preserved in mild to moderate AD and HD, whereas mental rotation is impaired. In part, this could be attributable to strategy choices made by patients.

Millions of people nationwide suffer from neurodegenerative disorders. Two such neurological conditions are Huntington’s disease (HD) and Alzheimer’s disease
(AD). HD is an inherited disease characterized by the degeneration of nerve cells in the basal ganglia and cerebral cortex that hinders function, movement, and cognitive abilities (Labuschagne et al., 2016). AD is a neurological disorder involving brain atrophy in the frontal cortex and hippocampus that leads to declines in thinking, behavior, and social abilities (Salimi et al., 2019). Because these two disorders involve degeneration in distinct parts of the brain, the resulting cognitive deficits differ. The purpose of the current study is to examine differences between these two patient groups in their visuospatial and mental rotation abilities.

Although HD is often considered a disorder of motor movements, cognitive impairments and psychiatric symptoms (e.g., irritability, depression, paranoia, delusions, and hallucinations) are also common in patients with HD. These deficits are wide-ranging, but recent studies have found that HD can lead to visuospatial impairments. For instance, one study by Labuschagne et al. (2016) evaluated differences in map search and mental rotation of a 3D cube between patients with HD and normal controls. Their results indicated that individuals with HD performed more poorly in comparison to the control group on both tests (Labuschagne et al., 2016). An additional study, by Majerová et al. (2011), compared spatial navigation abilities and cognitive abilities in people with HD and normal controls. The researchers concluded that in addition to displaying executive dysfunction and motor impairments, HD patients performed more poorly than normal controls on the spatial navigation assessment, and those spatial navigation abilities were directly correlated with disease severity, manifesting only in patients in moderate stages of the disease (Majerová et al., 2011). Another study, by Corey-Bloom et al. (2016), examined how HD patients compared to normal controls with regard to their basic visual perceptual abilities. Their results showed that HD patients performed worse than normal controls and that performance levels were again related to disease severity, although impairments were observed even in patients with mild HD (Corey-Bloom et al. 2016). It is evident that visuospatial function is compromised because of HD and gets worse as the disease progresses.

Various studies have also linked visuospatial deficits to the progression of AD. One study by Quental et al. (2013) examined the visuospatial abilities of AD patients compared to normal controls using an eight-test battery. Their findings indicated a significant difference between the control group and AD patients on six of the eight tests (Quental et al., 2013). Another study, by Salimi et al. (2019), similarly concluded that patients with AD were impaired on several visuospatial tasks compared to a control group. Beyond documenting deficits in the visuospatial abilities of patients with AD, past research has also demonstrated that visuospatial abilities predict everyday functioning in activities of daily living and may be a strong indicator of Alzheimer’s disease severity (Perry & Hodges, 2000). Together, these results point to the importance of understanding visuospatial abilities in patients with neurodegenerative disorders.
Although visuospatial deficits emerge from the pathophysiology of both HD and AD, few studies have compared the visuospatial abilities of individuals with HD and individuals with AD to one another. One such study was done by Lineweaver et al. (2005), who compared the ability of patients with AD, patients with HD, older healthy controls, and younger healthy controls to mentally rotate a stick figure. Although HD patients and AD patients both performed more poorly in mental rotation compared to their controls, they struggled with different aspects of the task. Response time for HD patients increased more relative to controls as the angle of orientation increased, highlighting that patients with HD have impaired information-processing speed. For AD patients, in contrast, accuracy decreased as the angle of orientation increased, indicating deficits in AD patients’ ability to mentally rotate the figure (Lineweaver et al., 2005). Another study, by Rouleau et al. (1992), compared the performance of HD patients, AD patients, and healthy controls on the clock-drawing test. They determined that patients with HD and patients with AD had deficits in drawing clocks compared to a control group but, again, made different mistakes. HD patients exhibited deficits in visuospatial planning and tended to draw small clocks. In contrast, AD patients showed impairments in visuospatial layout and drew large clocks with the numbers on the outside (Rouleau et al., 1992). These studies indicate that there are differences in the visuospatial deficits associated with HD and AD, but the extent of these differences is not currently known, because of the limited number of studies that have compared these two patient groups.

The current study aims to build on the previous research of Lineweaver et al. (2005) and Rouleau et al. (1992) by directly comparing the visuospatial abilities of HD patients, AD patients, and their respective control groups using a wider array of visuospatial tests that have both nonrotational and rotational components. Understanding the similarities and differences in visuospatial and mental rotation abilities in HD patients and AD patients is important because it could provide insight into disease progression as well as help differentiate brain regions implicated by each disease (Labuschagne et al., 2016; Perry & Hodges, 2000). The first goal of this study is to determine whether there are differences in visuospatial abilities between HD patients and younger healthy controls as well as between AD patients and older healthy controls. The second goal of this study is to compare the visuospatial deficits of patients with HD and patients with AD. I hypothesize that there will be significant differences in visuospatial capabilities between patient groups and their respective control groups. I also expect the pattern of visuospatial impairments to vary across patients with HD and patients with AD.
Method

Participants

This study included 18 HD patients (\(M_{\text{age}} = 42.39, SD = 8.54\)) and 20 young healthy controls (YHC) that were similar in age (\(M_{\text{age}} = 43.55, SD = 9.20\)) to the HD patients (\(p = .638\)). HD patients were diagnosed by a neurologist through family history, the existence of choreiform movements, and the presence of dementia. An additional 18 AD patients (\(M_{\text{age}} = 75.83, SD = 7.19\)) were recruited from the University of California San Diego Alzheimer’s Disease Research Center. Twenty older healthy controls (OHC) were matched with the AD patients on age, \(M_{\text{age}} = 76.95, SD = 4.48, p = .650\). Both YHCs and OHCs were recruited from healthy participants who were already enrolled in active research studies at the Alzheimer’s Disease Research Center or the San Diego VA Medical Center. All four groups were statistically similar in their years of education (\(F(3, 82) = 1.16, p = .383, \eta^2_p = 0.041\)) and gender distribution, \(\chi^2(n = 76) = 0.29, p = .81\).

Materials

Tests of Dementia Severity and Visual Perceptual Skills

Dementia Rating Scale (Mattis, 1988)

The Dementia Rating Scale (DRS) measures changes in cognition that result from dementia. It has five subsections, which assess different aspects of dementia: Attention, Initiation and Preservation, Construction, Conceptualization, and Memory.

Visual Reproduction Figure Matching (Wechsler, 1987)

Cards, each containing a single figure from the Visual Reproduction subtest of the Wechsler Memory Scale-Revised, were presented alongside four similar figures on a piece of paper. After examining each figure, participants identified which figure on the paper matched the one on the card.

Judgment of Line Orientation (Benton et al., 1983)

In the Judgment of Line Orientation Test (JLOT), participants were asked to determine the orientation of lines by matching the rotational degree of the lines to radii that were evenly spaced on a semicircle. Although there are typically 30 items in the test, only the odd-numbered items were used in this study.
Tests of Mental Rotation

Right-Left Orientation Test (Benton et al., 1983)

The first section of the Right-Left Orientation test (RLOT) included 12 items in which participants were asked to identify lateralized body parts on their own body (no rotation condition). In the second section, there were 8 items in which participants identified lateralized body parts on the examiner’s body while sitting face-to-face. When identifying the examiner’s left arm, the participant had to determine which was the examiner’s left versus right (rotation condition). Participants earned one point for each correct item, and the two sections were scored separately.

Luria Mental Rotation Test (Golden et al., 1980)

The Luria test has 10 items. Each item presents one target box that has a black border on the bottom and a circle in one corner. The boxes are shown in different orientations, and the participant has to mentally rotate the target box to identify a match to one of four other box options with different arrangements of borders and circles. Choices were presented simultaneously, and one point was awarded for each correct answer.

Money Road Map Test (Money, 1976)

The Money Road Map Test (MRMT) asks participants to determine the direction of turns on a route to assess right-left orientation. A path is traced on a map and each participant is asked if they would turn left or right at each change of direction if they were walking on the path. The path moves away from the participant (unrotated turns), perpendicular to the participant (90° rotated turns), and toward the participant (180° rotated turns). In this test, there are 32 changes of direction. The first half of the test includes seven unrotated turns, six 90° rotated turns, and three 180° rotated turns. The second half of the test contains two unrotated turns, seven 90° rotated turns, and seven 180° rotated turns. The total number of correct responses are recorded separately for changes of direction requiring no rotation, 90° rotation, and 180° rotation.

Stick Construction Test (Benson & Barton, 1970; Butters & Barton, 1970)

The Stick Construction Test (SCT) has two parts consisting of 10 items each in which participants were given five popsicle sticks that each had one tip blackened. In the first part, the examiner created a pattern out of sticks, which the participant had to duplicate using their own sticks while sitting side-by-side with the examiner.
(matching, no rotation, condition). The objective of the second part was the same—to copy the examiner’s pattern—except this time, the participant was sitting across the table from the examiner and had to duplicate the pattern from a flipped perspective (rotation condition). The two parts were scored separately, and one point was awarded for each correctly completed pattern.

**Evaluation of Strategy Use on the Mental Rotation Measures**

Upon completion of each individual test of mental rotation (Right-Left Orientation test, Luria test, Money Road Map Test, and Stick Construction Test), participants were given a list of strategies and were instructed to rank order all of the strategies they utilized while completing the test. These strategies were then classified by the researchers into personal rotation, extrapersonal rotation, or other. In personal rotation, someone mentally rotates themselves to a different spatial position to complete the task. In extrapersonal rotation, someone mentally rotates the object involved in the test to a different spatial position to complete the task. A strategy designated as “other” does not involve any mental rotation but can include logic and reasoning, such as deciding which side something is on and choosing the opposite rather than rotating something 180°. Based on the strategy that participants reported using the most, they were grouped into those who primarily used personal rotation, extrapersonal rotation, or an “other” strategy for each individual test. A participant could be in the personal mental rotation group for one test and an extrapersonal rotation group for a different test, meaning each participant ended up in a particular strategy group on each of the individual tests.

**Procedure**

Each participant in the study was tested individually after giving informed consent. All participants provided basic demographic information before completing the DRS. Next, researchers administered the RLOT, the visual reproduction test, the MRMT, the JLOT, the Stick Construction Test, and the Luria Mental Rotation Test in a fixed order.

**Data Analysis**

To analyze my data, I began by comparing the four groups on the visual reproduction test and JLOT using a one-way analysis of variance to determine if there were group differences in basic visual perceptual skills. I then analyzed the remaining tests to examine how each group performed on tests of mental rotation. For the Luria test, I used a one-way analysis of variance with Group as the between-subjects variable. For the RLOT, MRMT, and Stick Construction Test (all of which involve two or three different degrees of mental rotation), I ran a series of two-way analyses of
variance with Degree of Rotation as a within-subjects variable and Group as a between-subjects variable. To further examine interaction effects, I used simple main effects analyses with a Bonferroni correction to control the type I error rate. To further examine any significant group differences, I utilized follow-up analyses with Tukey’s post-hoc tests.

Next, I explored the strategies that participants used and how these influenced their performance on the tests of mental rotation. First, based on the categorization of participants into strategy groups based on their approach to each mental rotation test, I used a series of chi-square tests to compare the strategies most commonly utilized by participants from each of the four participant groups. For the mental rotation tests where the distribution of strategies was unequal across groups (MRMT, Stick Construction Test), I ran a series of follow-up chi-square analyses that compared the strategy choices of (1) the two healthy control groups, (2) each of the patient groups to their respective control groups, and (3) the two patient groups to each other (Bonferroni-corrected p-value = .05 / 3 = .017). I then identified the strategy that was most commonly utilized by participants in the healthy control groups and designated this the “typical HC strategy.” Next, I used a two-way analysis of variance to compare the performance of healthy controls who used the typical strategy to those who did not on the MRMT and Stick Construction Test. Degree of Rotation was a within-subject variable, and Strategy Group (typical HC vs. atypical HC) was a between-subject variable in these analyses. Finally, I examined the performance of patient groups who used the typical HC strategy and those who did not using another two-way Degree of Rotation by Strategy Group analysis of variance.

Results

Group Differences in Visual Perceptual Skills

Scores of each group on the Visual Reproduction test and the JLOT, the two tests of visual perceptual skills, are summarized in Table 1.
Table 1. Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>YHC (n = 20)</th>
<th>HD (n = 18)</th>
<th>OHC (n = 20)</th>
<th>AD (n = 18)</th>
</tr>
</thead>
</table>
| Age*             | 43.55 (9.20)
|                  | 42.39 (8.54)
|                  | 76.95 (4.48)
|                  | 75.83 (7.19) |
| Years of education| 14.48 (2.02) | 13.72 (2.24) | 15.25 (2.97) | 14.17 (3.60) |
| Gender           | 11 Female    | 8 Female    | 12 Female    | 10 Female   |
|                  | 9 Male       | 10 Male     | 8 Male       | 8 Male      |
| Race*            | 5 Black or   | 0 Black or  | 0 Black or   |
|                  | African      | African     | African      |
|                  | 1 Hispanic   | 0 Hispanic  | 1 Hispanic   |
|                  | 14 White     | 18 White    | 19 White     |
| Handedness       | 18 RH        | 17 RH       | 17 RH        |
|                  | 2 LH         | 1 LH        | 3 LH         |
|                  |              |             |              |
| DRS*             | 141.30 (1.78) | 130.56 (8.50) | 140.40 (3.72) | 120.56 (6.46) |
| Apraxia          | 59.95 (0.22) | 59.89 (0.32) | 59.95 (0.22) | 59.94 (0.24) |
| Visual reproduction* | 4.65 (0.67) | 3.78 (0.94) | 4.55 (0.51) | 3.94 (0.87) |
| JLOT*            | 13.25 (2.49) | 9.33 (3.79) | 12.65 (1.81) | 10.06 (2.36) |

Note: Patient groups with different superscripts differ from each other in their descriptive statistics (p < .05). AD = Alzheimer’s disease group; DRS = Dementia Rating Scale; HD = Huntington’s disease group; JLOT = Judgment of Line Orientation Test; LH = left-handed; OHC = older healthy control group; RH = right-handed; YHC = young healthy control group.

*Group distributions differ p < .05.

In the one-way analysis of variance examining scores on the Visual Reproduction test, a significant main effect of group emerged, $F(3, 72) = 6.18, p = .001, \eta_p^2 = 0.21$. Follow-up analyses indicated that the YHCs and OHCs performed similarly on this test ($p = .678$), but each healthy control group significantly outperformed their respective patient group (YHC vs. HD: $p = .001$; OHC vs. AD: $p = .006$). Scores of the two patient groups did not differ from each other on this test ($p = .513$).

The JLOT results paralleled those from the Visual Reproduction test, with a significant main effect of group, $F(3, 71) = 9.55, p < .001, \eta_p^2 = 0.29$. Again, YHCs and OHCs earned similar scores ($p = .483$), but both patient groups earned lower scores than their respective healthy control group (YHC vs HD: $p < .001$; OHC vs. AD: $p = .005$).
HD patients did not significantly differ from AD patients in their JLOT scores ($p = .0428$).

**Group Differences in Mental Rotation**

For the RLOT, a 2 (Rotation: $0^\circ$, $180^\circ$) x 4 (Group: YHC, HD, OHC, AD) analysis of variance resulted in a significant main effect of rotation ($F(1, 72) = 502.63$, $p < .001$, $\eta_p^2 = 0.88$) and a significant main effect of group, ($F(3, 72) = 5.96$, $p = .001$, $\eta_p^2 = 0.20$); however, these main effects were qualified by a significant two-way Rotation by Group interaction, $F(3, 72) = 4.52$, $p = .006$, $\eta_p^2 = 0.16$ (Figure 1). To further examine this interaction, we ran separate univariate analyses of variance for how the four groups performed at $0^\circ$ and at $180^\circ$ rotation. We applied a Bonferroni correction to the critical $p$-value to protect against a Type I error ($p = .05 / 2 = .025$). At the $0^\circ$ rotation, the main effect of Group did not reach significance, $F(3, 72) = 1.68$, $p = .18$, $\eta_p^2 = 0.07$. At $180^\circ$ rotation, the main effect of Group did reach significance, $F(3, 72) = 5.69$, $p = .001$, $\eta_p^2 = 0.19$. Follow-up analyses indicated that the two healthy control groups performed similarly at $180^\circ$ rotation ($p = .882$). Both patient groups performed more poorly than their respective healthy control groups (YHC vs. HD: $p = .002$; OHC vs. AD: $p = .015$), but the two patient groups did not significantly differ from each other in their performance ($p = .532$).
Figure 1. Performance of Each Group on the Right-Left Orientation Test

Note. A significant interaction emerged between Group and Degree of Rotation. The groups performed equivalently at 0° rotation but differed at 180° rotation. Mean differences between groups are denoted by different superscripts. Both patient groups performed significantly worse than their relative control groups when required to rotate 180°.

AD = Alzheimer’s disease group; HD = Huntington’s disease group; OHC = older healthy control group; YHC = young healthy control group.

The results of a one-way analysis of variance examining scores on the Luria test revealed a significant main effect of Group, $F(3, 72) = 6.60, p = .001, \eta_p^2 = 0.22$ (Figure 2). The YHCs and OHCs did not differ from each other on this test ($p = .685$), but both of the patient groups performed significantly more poorly than their respective control groups (YHC vs. HD: $p = .001$; OHC vs. AD: $p = .004$). The two patient groups performed identically on the Luria test ($p = 1.00$).
Figure 2. Performance of Each Group on the Luria Test

Note. A significant main effect of Group emerged. Mean differences between groups are denoted by different superscripts. Both patient groups performed significantly worse than their relative control groups on this measure of mental rotation.

AD = Alzheimer's disease group; HD = Huntington's disease group; OHC = older healthy control group; YHC = young healthy control group.

Scores on the MRMT were analyzed using a 3 (Rotation: 0°, 90°, 180°) x 4 (Group: YHC, HD, OHC, AD) analysis of variance, which resulted in a significant main effect of Rotation ($F(1, 72) = 136.07, p < .001, \eta^2_p = 0.65$) and a significant main effect of Group, $F(3, 72) = 26.00, p < .001, \eta^2_p = 0.27$. These main effects were again qualified by a significant two-way Rotation by Group interaction, $F(3, 72) = 7.00, p < .001, \eta^2_p = 0.27$ (Figure 3). Separate univariate analyses of variance comparing each group’s performance at 0°, 90°, and 180° rotation more closely examined this interaction. We applied a Bonferroni correction to the critical $p$-value to protect against a Type I error ($p = .05 / 3 = .017$). At the 0° rotation, the main effect of Group did not reach significance, $F(3, 72) = 1.97, p = .13, \eta^2_p = 0.08$. At both the 90° rotation ($F(3, 72) = 6.93, p < .001, \eta^2_p = 0.22$) and the 180° rotation ($F(3, 72) = 4.72, p = .005$), the main effect of group did reach significance. The two healthy control groups performed similarly at the 90° ($p = .32$) and 180° rotation ($p = .19$). Both patient groups performed more poorly than their respective healthy control groups at the 90° rotation (YHC vs. HD: $p = .004$; OHC vs. AD: $p = .003$), but only HD patients performed more poorly than their healthy
controls at the 180° rotation (YHC vs. HD: $p = .004$; OHC vs. AD: $p = .054$). The two patient groups did not significantly differ from each other on their performance at the 90° rotation ($p = .29$) or the 180° rotation ($p = .80$).

Figure 3. Performance of Each Group on the Money Road Map Test

![Figure 3](image)

Note. A significant interaction emerged between Group and Degree of Rotation. The groups performed equivalently at 0° rotation but differed at both 90° and 180° rotation. Mean differences between groups are denoted by different superscripts. Both patient groups performed significantly worse than their relative control groups when required to rotate 90°, but only the HD group performed significantly worse than their controls at 180° rotation.

AD = Alzheimer’s disease group; HD = Huntington’s disease group; OHC = older healthy control group; YHC = young healthy control group.

A 2 (Rotation: 0°, 180°) x 4 (Group: YHC, HD, OHC, AD) analysis of variance on the Stick Construction Test resulted in a significant main effect of Rotation ($F(1, 72) = 136.07, p < .001, \eta^2_p = 0.65$) and a significant main effect of Group, $F(3, 72) = 19.42, p < .001, \eta^2_p = 0.45$. These main effects were again qualified by a significant two-way Rotation by Group interaction, $F(3, 72) = 17.20, p < .001, \eta^2_p = 0.42$ (Figure 4). To further examine this interaction, we ran separate univariate analyses of variance comparing each group’s performance at 0° and 180° rotation. We applied a Bonferroni correction to the critical $p$-value to protect against a Type I error ($p = .05 / 2 = .025$).
the 0° rotation, the main effect of Group did not reach significance, $F(3, 72) = 2.83, p = .04, \eta_p^2 = 0.11$. At 180° rotation, the main effect of Group was significant, $F(3, 72) = 19.32, p < .001, \eta_p^2 = 0.45$. The two healthy control groups performed similarly ($p = .27$). Both patient groups performed more poorly than their respective healthy control groups (YHC vs. HD: $p < .001$; OHC vs. AD: $p < .001$). The two patient groups did not significantly differ from each other on their performance on the Stick Construction Test at 180° rotation ($p = .04$).

**Figure 4. Performance of Each Group on the Stick Construction Test**

![Graph showing performance of each group on the Stick Construction Test]

**Note.** A significant interaction emerged between Group and Degree of Rotation. The groups performed equivalently at 0° rotation but differed at 180° rotation. Mean differences between groups are denoted by different superscripts. Both patient groups performed significantly worse than their relative control groups when required to rotate 180°.

AD = Alzheimer’s disease group; HD = Huntington’s disease group; OHC = older healthy control group; YHC = young healthy control group.

**Group Differences in Strategy Use**

A chi-square test examining the proportion of participants who used a personal mental rotation strategy, an extrapersonal rotation strategy, or an “other” strategy on each of the mental rotation tests revealed similar distributions of strategy choice across the four participant groups on both the RLOT ($\chi^2(n = 75, df = 6) = 3.14, p = .792$) and the Luria test, $\chi^2(n = 73, df = 6) = 4.43, p = .619$ (Table 2). Significant differences did emerge in how commonly the four participant groups used each of the three strategies
on the MRMT ($\chi^2(n = 73, df = 6) = 12.91, p = .044$) and the Stick Construction Test, $\chi^2(n = 68, df = 6) = 15.55, p = .016$. Follow-up analyses (Bonferroni-corrected $p$-value = .5 / 3 = .018) indicated that the YHCs and OHCs did not differ in their strategic approaches on either the MRMT ($\chi^2(n = 40, df = 1) = 0.23, p = .633$) or Stick Construction Test, $\chi^2(n = 39, df = 2) = 3.29, p = .193$. A comparison of HD patients to YHCs revealed a different pattern of strategy selection on the MRMT ($\chi^2(n = 40, df = 1) = 8.26, p = .016$), but not the Stick Construction Test, $\chi^2(n = 35, df = 2) = 5.87, p = .053$. AD patients did not differ significantly from the OHCs on either the MRMT ($\chi^2(n = 36, df = 1) = 2.40, p = .121$) or Stick Construction Test, ($\chi^2(n = 33, df = 2) = 6.55, p = .038$) with the Bonferroni correction applied. Finally, no significant differences in strategy choice emerged when comparing HD patients to AD patients on either test (MRMT: $\chi^2(n = 33, df = 2) = 1.48, p = .477$; Stick Construction Test: $\chi^2(n = 29, df = 2) = 2.56, p = .324$).

Table 2. Self-Reported Frequency of Use of Strategy Type by Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Personal rotation strategy n (%)</th>
<th>Extrapersonal rotation strategy n (%)</th>
<th>Other strategy n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Luria test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHC</td>
<td>2 (10.0)</td>
<td>14 (70.0)</td>
<td>4 (20.0)</td>
</tr>
<tr>
<td>HD</td>
<td>2 (13.3)</td>
<td>10 (66.7)</td>
<td>3 (20.0)</td>
</tr>
<tr>
<td>OHC</td>
<td>1 (5.0)</td>
<td>16 (80.0)</td>
<td>3 (15.0)</td>
</tr>
<tr>
<td>AD</td>
<td>0</td>
<td>12 (66.7)</td>
<td>6 (33.3)</td>
</tr>
<tr>
<td></td>
<td>RLOT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHC</td>
<td>6 (31.6)</td>
<td>0</td>
<td>13 (68.4)</td>
</tr>
<tr>
<td>HD</td>
<td>5 (27.8)</td>
<td>0</td>
<td>13 (72.2)</td>
</tr>
<tr>
<td>OHC</td>
<td>6 (30.0)</td>
<td>1 (5.0)</td>
<td>13 (65)</td>
</tr>
<tr>
<td>AD</td>
<td>3 (16.7)</td>
<td>1 (5.6)</td>
<td>14 (77.8)</td>
</tr>
<tr>
<td></td>
<td>MRMT*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHC</td>
<td>18 (90)</td>
<td>0</td>
<td>2 (10.0)</td>
</tr>
<tr>
<td>HD</td>
<td>8 (47.1)</td>
<td>1 (5.9)</td>
<td>8 (47.1)</td>
</tr>
<tr>
<td>OHC</td>
<td>17 (85.0)</td>
<td>0</td>
<td>3 (15.0)</td>
</tr>
<tr>
<td>AD</td>
<td>10 (62.5)</td>
<td>0</td>
<td>6 (37.5)</td>
</tr>
<tr>
<td></td>
<td>Stick Construction Test*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YHC</td>
<td>10 (50.0)</td>
<td>3 (15.0)</td>
<td>7 (35.0)</td>
</tr>
<tr>
<td>HD</td>
<td>4 (26.7)</td>
<td>0</td>
<td>11 (73.3)</td>
</tr>
<tr>
<td>OHC</td>
<td>13 (68.4)</td>
<td>4 (21.1)</td>
<td>2 (10.5)</td>
</tr>
<tr>
<td>AD</td>
<td>6 (42.9)</td>
<td>1 (7.1)</td>
<td>7 (50)</td>
</tr>
</tbody>
</table>

Note. AD = Alzheimer’s disease group; HD = Huntington’s disease group; MRMT = Money Road Map Test; OHC = older healthy control group; RLOT = Right-Left Orientation test; YHC = young healthy control group. *Group distributions differ $p < .05$. 

243
Performance on Tests of Mental Rotation Given Strategy Use

Two sets of analyses of variance examined how strategy choice affected performance on the MRMT and Stick Construction Test. First, I compared healthy controls who used the typical strategies with healthy controls who used an alternate strategy. Second, I evaluated how patients who used the typical healthy controls (HC) strategy performed relative to those who used an alternate strategy. I will focus the discussion on the Strategy main effect and the Strategy by Rotation interaction because the Rotation main effect was addressed previously.

For the MRMT, the 3 (Rotation: $0^\circ$, $90^\circ$, $180^\circ$) x 2 (Strategy Group: Typical HC, Atypical HC) analysis of variance based on HC data showed no main effect of strategy choice ($F(1, 38) = 2.73, p = .107, \eta_p^2 = 0.067$), and strategy choice did not interact with rotation to affect performance, $F(2, 37) = 2.84, p = .071, \eta_p^2 = 0.133$ (Figure 5, panel A). For patients, strategy choice did not interact with rotation to affect performance, ($F(2, 30) = 0.82, p = .449, \eta_p^2 = 0.052$), but patients who used the typical HC strategy significantly outperformed those who used an alternate approach on the MRMT, $F(1, 31) = 4.79, p = .036, \eta_p^2 = 0.134$ (Figure 5, panel B).
Figure 5. Healthy Controls’ and Patients’ Performance on the Money Road Map Test by Strategy

Note. The typical healthy control (HC) strategy involved personal mental rotation, while an atypical HC strategy involved a different tactic such as logic or extrapersonal rotation. Neither the mean effect of Strategy Choice nor the Strategy Choice x Rotation interaction reached significance for the healthy controls. Patients performed significantly better on the test overall when they utilized the typical HC strategy than when they used an alternate approach, regardless of the degree of rotation.

For the Stick Construction Test, the 2 (Rotation: 0°, 180°) x 2 (Strategy Group: Typical HC, Atypical HC) analysis of variance based on HC data showed no main effect of strategy choice ($F(1, 37) = 0.08, p = .776, \eta^2_p = 0.002$), and strategy choice did not interact with rotation to affect performance, $F(1, 37) = 0.03, p = .873, \eta^2_p = 0.001$ (Figure 6, panel A). For patients, there was also no main effect of strategy choice ($F(1,
$27) = 1.95, p = .174, \eta^2_p = 0.067$) or interaction between strategy choice and rotation, $F(1, 27) = 0.64, p = .431, \eta^2_p = 0.023$ (Figure 6, panel B).

Figure 6. Healthy Controls’ and Patients’ Performance on the Stick Construction Test by Strategy

Note. The typical healthy control (HC) strategy involved personal mental rotation, while an atypical HC strategy involved a different tactic such as logic or extrapersonal rotation. Neither the mean effect of Strategy Choice nor the Strategy Choice x Rotation interaction reached significance for either group.
Discussion

The purpose of this study was to directly compare the visuospatial abilities of HD patients, AD patients, and their respective controls by using visuospatial tests with and without a mental rotation component. The first goal of this study was to determine whether differences in mental rotation abilities emerged between HD patients and YHCs and between AD patients and OHCs. The second goal of this study was to compare HD patients and AD patients in their performance of mental rotation. I hypothesized that patient groups would perform significantly worse on the tests of mental rotation compared to healthy control groups and that HD patients and AD patients would have different patterns of impairment across the visuospatial tests. The results indicate that both patient groups performed significantly worse relative to their respective controls on nearly every visuospatial test. Furthermore, the visuospatial abilities of the HD patients and the AD patients did not differ from each other.

Compared to age-matched healthy controls, HD patients and AD patients have many commonalities in their visuospatial deficits. Both HD patients and AD patients struggled on the Visual Reproduction test and JLOT, but neither group was impaired on the nonrotational components of the mental rotation tests. Although these results may seem contradictory, it is likely that the nonrotational components on the mental rotation tests were easier than the tests measuring visuoperceptual skills, which are designed to be sensitive to impairments in visuospatial abilities. On every mental rotation test—RLOT, Luria test, MRMT, and Stick Construction Test—both patient groups performed significantly worse than their respective healthy control group on the rotation components; thus, both patients with HD and patients with AD showed deficits in their mental rotation abilities. The finding that patient groups perform worse than controls on most visuospatial tests included in this study is not surprising, as other research has concluded that HD patients (Corey-Bloom et al., 2016) and AD patients (Quental et al. 2013; Salimi et al. 2019) exhibit deficits on basic perceptual abilities and that HD patients have more deficits in rotational abilities compared to controls (Labuschagne et al., 2016). Only one distinct pattern emerged in the comparisons between HD patients, AD patients, and their control groups. HD patients were significantly impaired relative to the YHCs on the 180° rotation of the MRMT, whereas AD patients were not significantly impaired when compared to the OHCs. Further research would be necessary to replicate this effect and determine what may be driving this difference.

In direct comparison of HD patients with AD patients on all the tests of mental rotation (RLOT, Luria test, MRMT, and Stick Construction Test), there were no significant differences in performance between the two groups on either nonrotational or rotational components. Although only a few studies have directly compared visuospatial deficits of HD patients and AD patients, the current findings
are inconsistent with results from Lineweaver et al. (2005) and Rouleau et al. (1992) that HD patients and AD patients are impaired on distinct aspects of mental rotation. Lineweaver et al. (2005) found that HD patients had a greater response time when compared to controls but maintained accuracy whereas AD patients had decreased accuracy but similar reaction times relative to controls. This indicates that HD patients have deficits related to information-processing speed, while AD patients have deficits with mental rotation. Perhaps I failed to find differences between the two groups because all tests in this study were untimed. Similarly, Rouleau et al. (1992) found different patterns of clock drawing between HD and AD patients. HD patients drew small clocks relative to their control group, whereas AD patients drew larger clocks with the numbers outside the clock. This suggested that HD patients have deficits with visuospatial planning and AD patients have deficits with visuospatial layout. These planning deficits in HD patients documented by Rouleau et al. (1992) may reflect the executive dysfunction associated with HD, which might have contributed to the differences in strategy use I found across groups.

Participants varied in the strategies they used to complete tests of mental rotation. On some tests, they tended to use an extrapersonal rotation strategy (Luria test), whereas on others, they tended to use a personal rotation strategy (MRMT, Stick Construction Test). The results of the Stick Construction Test surprised me because I thought it would be easier to mentally rotate the sticks rather than oneself to complete this task. Group differences emerged in strategy selection on both the MRMT and the Stick Construction Test. On both of these tests, HD patients were much more likely to use an “other” strategy than were their healthy controls, who were more likely to use a personal rotation strategy. Because of low power in these analyses (many patients were unable to report their strategies), this difference reached significance only for the MRMT. AD patients showed a similar pattern only on the Stick Construction Test, but this difference did not reach significance in follow-up analyses.

The finding that the HD patients utilized different strategies than the YHCs did on the MRMT could have in part contributed to their poor performance on this test. Patients who used the common HC strategy when completing the MRMT performed significantly better compared to patients who used an alternate strategy; thus, the HD patients’ tendency to not use a personal rotation strategy, in comparison to the AD patients’ more common use of a personal rotation strategy, on this test may help explain why HD patients were impaired relative to their controls on the 180° rotation of the MRMT while AD patients were not.

Limitations and Future Directions

There are some limitations to this study. First, the study included a homogenous sample of healthy controls and patients who were mostly White and
highly educated. A more diverse sample may yield different patterns of results. Second, patients in this study had only mild to moderate dementia as measured by their DRS scores. Perhaps as dementia progresses, additional or more differential patterns of deficits may emerge between HD patients and AD patients. Additionally, patients in the study had mild difficulties with the Visual Reproduction test and JLOT. These two tests measure basic visual abilities, which could have added to patients’ difficulties when having to perform tasks of mental rotation. At the same time, the lack of significant deficits shown by HD and AD patients on the nonrotational components of the mental rotation tests suggests they had the basic skills needed to perform these tasks. Furthermore, the analysis of strategies is based off of self-reports, and there is no way to determine if participants were accurate in sharing their approaches at the completion of each test. This is especially true for the patients who were diagnosed with dementia, many of whom were unable to respond to the question about their strategy use, limiting the number of participants included in the strategy analyses. Additionally, strategies were not presented distinctly as choices of personal mental rotation, extrapersonal mental rotation, or other; rather, some categories of strategies were presented through multiple examples whereas others were not, and we deciphered which strategies belonged to which group in a post hoc fashion.

Conclusions

This study found that HD patients and AD patients have similar visuoperceptual and mental rotation deficits but have relatively spared nonrotational abilities. HD, but not AD, patients may use different strategies than their age-matched peers on some mental rotation tests, and this choice of strategic approach may contribute to their deficits. Future research should further explore the mental rotation abilities of these two patient groups using larger and more diverse samples. Gaining an increased understanding of visuospatial differences in HD patients and AD patients can give rise to knowledge about the progression of the diseases and the brain regions that underlie them.
References


DISCRIMINATION BY ASSOCIATION: THE ROLE OF RACE AND EXONERATION IN HIRING DECISIONS

EMMA BERGQUIST, BUTLER UNIVERSITY
MENTOR: FABIANA ALCESTE

Abstract

People convicted of felonies are often discriminated against when applying for jobs after prison (Mikkelson & Schweitzer, 2019). Interestingly, exonerated individuals are also discriminated against upon their release, even though they were wrongfully convicted. To date, there has been no direct comparison of the discrimination between these two groups when they are applying for jobs. The present study closes this gap by comparing the application strength of candidates with varying categories of race, legal status, and community support. We collected data from 299 CloudResearch participants online and presented them with job application materials that they evaluated as if they were hiring managers. Results showed that across all independent variables, the applicant who was guilty of the crime was rated significantly lower than the other candidates. Additionally, the applicant who falsely confessed was also viewed less favorably than the control group in terms of character. These findings suggest that attitude, policy, and program changes are needed to improve the reentry process for people who have been incarcerated, including those who have been exonerated.

After being released from incarceration, many individuals remain under the intense scrutiny of societal stigma that exists beyond prison walls. Society's negative perceptions of previously incarcerated individuals create roadblocks to the reentry process (Rade et al., 2016). Formerly incarcerated people often have difficulty obtaining housing, employment, and other necessities as they adjust to life after incarceration (Liem, 2016). As more accounts have begun to emerge of wrongful convictions being overturned, it is important to examine whether the stigma of prison colors perceptions of the wrongfully convicted as well. Although many average Americans may not believe they could be convicted of a crime they did not commit (Alceste et al., 2021; Henkel et al., 2008), data show that this phenomenon is occurring much more often than one may think. For instance, a recent news story broke that the tactics of a corrupt sergeant in Chicago have led to more than 200 overturned convictions (Mitchell, 2022). Stories of exonerees such as the Exonerated Five, five young minority teens who were wrongfully convicted of rape and assault in 1989
(Burns et al., 2012), illustrate the types of practices that may lead to individuals being convicted by juries of their peers, even when they have done nothing wrong. Cases such as these highlight the importance of studying the stigma and obstacles that exonerated individuals face after their official legal battles end.

After exonerees are released, they must reintegrate into a society that they left years, sometimes decades, before. When applying for a job, an exoneree may have to indicate that they have been convicted of a crime, with nowhere to explain that their conviction was wrongful and has since been overturned. The stigma attached to being incarcerated may lead to discrimination in many aspects of their new lives. Building managers, employers, and neighbors may not take kindly to knowing that a previously incarcerated individual is seeking their help (Liem, 2016). Building upon previous research, the present study asks how the legal status, race, and community support for an applicant affect their perceived competency, character, and application strength for a managerial job posting.

Wrongful Convictions Based on False Confession and Eyewitness Misidentification

The weight of societal bias against those involved in the criminal legal system is likely felt by all formerly incarcerated people. But perhaps the wrongfully convicted carry an additional burden when reentering a society that may not understand how they were convicted if they are innocent. The factors that contributed to their wrongful conviction might play an interesting role.

Two of the leading contributing factors for wrongful convictions are false confessions and mistaken eyewitness identifications (Innocence Project, 2023). Previous research shows that a false confession is often viewed as the fault of the convicted individual (Scherr et al., 2018, 2020). It is difficult for many people to imagine confessing to a crime that they had nothing to do with (Alceste et al., 2021; Henkel et al., 2008). Recent studies have found that most people view exonerees who falsely confessed to crimes as unintelligent or mentally ill (Scherr et al., 2018, 2020). This perception exists beyond the instance of the false confession as well, implying that the wrongfully convicted individual is to blame for their incarceration, and often, doubts remain about their true innocence, regardless of evidence on their behalf. In turn, this undermines people’s willingness to support reintegration processes for these exonerees when the people are not fully convinced of exonerees’ innocence (Pica et al., 2022; Scherr et al., 2018) Moreover, an examination of the time spent in prison for wrongfully convicted individuals by Scherr & Normile (2022) demonstrated that false confessions predict long delays (772 days on average) between release and official exoneration. This is almost a year longer than the average delay for a wrongfully convicted individual who did not falsely confess. In addition, judges are less willing to grant appeals on cases involving false confessions. This is likely due to a combination of the inability to understand the mechanisms that lead to a false
confession, the compelling nature of the confessions themselves, and confirmation bias rooted in society that results in much of the public continuing to believe that the false confessor is guilty (Scherr & Normile, 2022). A delay such as this takes time away from resources available to the exoneree, such as housing assistance and job training. Even after release, exonerees are not given access to reentry programming that other previously incarcerated individuals are afforded (Jackson et al., 2021).

Mistaken eyewitness testimony might not have the same prejudicial effect, however, perhaps because the public can better understand someone making an error during a lineup procedure, as opposed to someone confessing to a crime they did not commit. According to the Innocence Project (2023), approximately 69% of convictions overturned because of exculpatory DNA evidence were based on false eyewitness testimony. The reason for a wrongful conviction being outside of someone’s control, like false eyewitness testimony, may be something that laypeople understand could happen to any innocent person, which may welcome less judgment for exonerees who were convicted on factors beyond their control. One recent experiment highlighted this by asking participants to recommend compensation for exonerees after release from a wrongful conviction. The researchers found that individuals convicted based on false confessions were perceived to be less deserving of compensation than were exonerees convicted because of eyewitness testimony (Kieckhaefer & Luna, 2022). This further illustrates the importance of how an individual was wrongfully convicted and if that factor is perceived to be within the person’s control.

Studies have shown that less than half of the American public believes that those who have been to prison and are struggling to get jobs deserve the same resources as those who have not been to prison and are also searching for employment (Ouellette et al., 2017); however, researchers who interviewed formerly incarcerated individuals about the support they received upon their release found that positive social support for this group of people can be instrumental in their reentry journey (Kjellstrand et al., 2021). Additionally, in certain fields, letters of recommendation have been evaluated as one of the top three most important factors that an employer uses to make a hiring decision (Morgan & Landrum, 2012). In terms of employment, a recommendation letter on behalf of the applicant who has previously been incarcerated may increase the competitiveness of their application.

Another helpful form of support for formerly incarcerated individuals is funded reintegration programs, although one experiment found that the public viewed exonerees who falsely confessed as being less deserving of these programs and assistance than rightfully convicted people or those who did not confess to crimes and were then exonerated (Scherr et al., 2020). Overall, the present study examines the difference in ratings for applicants who have been wrongfully convicted based on false confessions and those who were wrongfully convicted based on eyewitness testimony.
to understand if this trend is consistent in a situation that formerly incarcerated people may find themselves in after release.

Conviction Prejudice

Though exonerated individuals have been cleared of any crime, such an individual must still indicate on a job application that they have been convicted of a felony. Ban the Box is a recent international movement focused on removing the section of employment applications that asks about the applicant’s criminal history (All of Us or None, n.d.). This movement is supported by research indicating that when equally qualified candidates are considered for the same position, the applicant without a criminal record is preferred over their counterpart (Graber & Zitek, 2022). Many activists have advocated for criminal disclosure to occur during the interview process or through a background check instead of on the initial job application. There may be a societal unconscious bias against people who have been tried and rightfully convicted in the United States, evidenced by the fact that mock jurors are more likely to convict a defendant with a previous conviction (Greene & Dodge, 1995).

Less studied is the stigma associated with those who have been wrongfully incarcerated. An interview-based study of suburban citizens’ perceptions of prison stigmatization found that most people believe that a stigma exists against exonerees yet claim they do not harbor biases themselves; however, the researchers found that most of the participants did harbor an unconscious bias against exonerated individuals because the individuals were affiliated with the criminal legal system at all (Blandisi et al., 2015). Specifically, this study along with an experiment from Scherr et al. (2020) found that, despite exonerees being cleared of the crimes they were convicted for, the public continues to stigmatize them because of their association with the criminal legal system. Recent evidence suggests that exonerees face housing discrimination when they are released from prison. Notably, exonerees applying for housing received the same response rate as those who had been rightfully convicted, much lower than applicants who had no affiliation with the criminal legal system (Kukucka et al., 2021). Additionally and importantly, though both exonerees and rightly convicted offenders may face employment discrimination, exonerees may face additional difficulties because they are not provided the same reintegration resources given to other people who were formerly incarcerated (Kukucka et al., 2019).

Additional discrimination may affect previously incarcerated individuals who are members of minority groups, specifically in terms of race. In her book *The New Jim Crow*, Michelle Alexander makes the argument that the criminal legal system has become a new way for the United States to discriminate against Black individuals (Alexander, 2011). This is true for both rightfully and wrongfully convicted Black people.
Race as a Contributing Factor

Racial discrimination is evident in multiple institutions in the United States, built upon hundreds of years of oppression and exploitation. Although much of society has made strides toward closing the racial disparity gap, racial discrimination is still a prevalent part of many American institutions. Experiments focused on the callback rates for applicants found that Black applicants were less likely to receive callbacks than were White applicants with the same credentials. Furthermore, criminal records lowered Black applicants’ callback rates but did not affect the callback rates of White applicants (DeWitt & Denver, 2020; Mobasser, 2019).

Racial discrimination already exists without the addition of a criminal record and thus worsens when a conviction is added, even if that conviction is baseless. Racial discrimination in employment comes primarily when White supervisors are in charge of the hiring decision (Avery et al., 2008). Considering that those in positions of power, who thus decide whom to hire particularly, have for decades been primarily White, this trend is particularly worrisome (Emeka, 2018; Wilson & Mossakowski, 2012). Across all fifty states, there is a disproportionately high rate of unemployment for Black individuals that does not exist on the same scale for White people (Emeka, 2018).

Previous literature also suggests that people of color are arrested at much higher rates than White people (D’Alessio & Stolzenberg, 2003). Research has found that 61% of those who have been released from prison through DNA exoneration are Black Americans (Howard, 2019). Though finding employment upon release as an exoneree of any race may be difficult, the race of the applicant can add further challenges if the applicant is not White.

Moreover, exoneration itself is less likely for people of color (Jokwi, 2017). In addition to being more likely to be falsely imprisoned, members of marginalized groups also spend more of their sentences in prison before being exonerated (Itskovich et al., 2023). As there is a proportionally much higher rate of incarceration for people of color (Blankenship et al., 2018), we hypothesize that society may see the wrongful incarceration of Black individuals as more justified than that of White individuals. Though each applicant in the present study stated the same reason for their original conviction (and exoneration, if relevant), the perception from the participant may be that the White applicant was always innocent while the Black applicant may have still committed the crime.

The present study was created to continue to examine whether the variable of race can lead to failed reentry and a return to prison for some exonerees. Specifically, Black individuals who have been incarcerated are seen as more aggressive and more likely to commit another crime than are their White counterparts (Howard, 2019). This is an example of stereotyping—a typically exaggerated negative generalization.
about a group of people even when these characteristics are not true of everyone within that group (American Psychological Association, 2022). This stereotype may play a role in allowing formerly incarcerated White individuals more employment opportunities.

The Present Study

Employee discrimination is the unequal treatment of a candidate based on an uncontrollable feature that the applicant possesses. For this study, this feature is a wrongful conviction and an exoneration for a crime that the applicant did not commit. Previous literature on this topic suggests that discrimination exists against rightfully convicted individuals, exonerees, and Black applicants. Though these variables have been examined separately, the present study investigates the interaction between these factors and the implications for individuals who have been falsely imprisoned. Specifically, we examine interactions between self-identified race, affiliation with the criminal legal system, reason for exoneration if applicable, and the presence of a recommendation letter.

Based on the previous research on this topic, the current experiment tests the following hypotheses:

H1: The applicant who is stated to be White, is unaffiliated with the criminal legal system, and has a recommendation letter will receive, on average, the highest rating across all four metrics from the participants.

H2: The applicants who are stated to be exonered and those who are stated to be rightfully convicted will receive similar ratings across all four metrics from the participants.

H3: The applicant who was wrongfully convicted based on eyewitness testimony will be preferred over the applicant who was convicted because of a coerced false confession, regardless of other factors.

Method

The current study is a 2 (race of applicant: Black or White) x 4 (legal status: no conviction, guilty, false confession, or mistaken eyewitness) x 2 (recommendation letter present or absent) between-subjects experimental design.

Participants

We recruited participants through CloudResearch (Litman et al., 2017), an online platform in which the general public can participate in research studies for compensation. We compensated each participant $1.50 for the session (median
duration approximately 12 minutes). We collected data from 413 participants. We excluded 102 participants’ data based on failed attention and manipulation checks: 11 participants failed the manipulation check for race, 62 incorrectly stated the legal status condition of the applicant, 27 did not correctly report the presence or absence of a recommendation letter, and 2 people attributed the recommendation letter to the wrong person. An additional 12 participant responses were discarded because the respondents did not complete the entirety of the survey. The final sample of participants (n = 299) was 50.8% male, 47.2% female, and 0.7% other; 1.3% preferred not to say. Additionally, 83.3% of participants were White, 8.4% were Black, 6.4% were Hispanic or Latinx, 1.7% were Native Hawaiian or Pacific Islander, and 0.3% preferred not to say. On average, participants were 41.8 years old (SD = 11). Additionally, 44.8% reported that they knew someone who had been to prison, and 46.6% had previous experience hiring others as part of their job.

Design and Materials

Race

The applicant in this study, named Michael, was either a Black man or a White man. Applicant race was displayed to participants in a demographic information section in the application, as well as a picture of Michael, randomly varied to be a White man or Black man from the Chicago Face Database (Ma et al., 2015), an open-access platform. We ran a pilot study with the pictures, and the subsequent paired t-test found no significant difference in the perceived attractiveness, \( t(49) = 0.99, p = .33, \text{Cohen’s } d = 0.14 \). There was also no significant difference in perceptions of how intimidating the person looked between the same two photographs of the men, \( t(49) = 0.82, p = .42, \text{Cohen’s } d = 0.12 \).

Legal Status

Participants were randomly assigned to one of four criminal-status conditions. Michael was either (1) rightfully convicted of a crime he committed, (2) wrongfully convicted based on mistaken eyewitness testimony, (3) wrongfully convicted based on a false confession and exonerated, or (4) not affiliated with the criminal legal system in any way. Concerning this variable, the application form contained information only about whether Michael had a criminal record. The specific reason for conviction or exoneration, if applicable, was explained in the applicant’s cover letter, which participants read after reviewing the initial application form. The cover letter for the guilty applicant stated that the applicant had been convicted of a crime he had committed and that he had been released after serving his full sentence. The exonerated applicant explained his reason for wrongful conviction in the cover letter, whether eyewitness misidentification or false confession, and maintained his
innocence. In all applications besides the control, the applicant had served four years in prison.

**Recommendation Letter**

Participants were randomly assigned to applications that either did or did not include a recommendation letter vouching for the applicant’s character. The letter came from Michael’s former teacher, who attributed positive qualities to the applicant. This letter was general and held constant throughout the conditions where it was applied. There was no mention of the applicant’s criminal history in the letter.

**Application Strength**

We asked participants about the strength of Michael’s application. This variable was measured by asking each participant to complete a 7-point Likert scale ranging from 1 (completely disagree) to 7 (completely agree) based on the statement “In my opinion, this was a strong application.” A higher score on this question indicates a stronger inclination in the participant to hire the person specified on the resume.

**Competency**

Participants also judged Michael’s competency to perform the job for which he applied. This was measured by asking each participant to complete a 7-point Likert scale ranging from 1 (completely disagree) to 7 (completely agree) in response to the statement “Based on what I know about this applicant, I think he is competent” and the reverse-coded statement “This applicant would require extensive training if he were hired.” We used the average score of the responses to these two questions to create a competency score. Responses to these two questions (using the straight-coded versions of the reversed questions) were sufficiently correlated (Cronbach’s $\alpha = .61$). A higher score on this composite indicates a higher level of perceived competency for the candidate based on his application materials.

**Character**

Lastly, we created a composite score measuring the participants’ overall judgment of Michael. Each participant completed a 7-point Likert scale ranging from 1 (completely disagree) to 7 (completely agree) in response to multiple statements. These statements were “This application suggests that the applicant is trustworthy,” “I would welcome this applicant in my workplace,” and two reverse-coded statements: “This applicant would require direct supervision if he were hired” and “I would feel uncomfortable working with this applicant.” The responses to these questions (using
the straight-coded versions of the reversed questions) were highly correlated with each other (Cronbach’s $\alpha = .91$).

**Manipulation Checks**

A series of manipulation and attention checks was included to ensure that the participants were actively engaged in the study. In the survey, we asked the participants to “Please select number five for this question” during the questions evaluating the dependent variables. Additionally, participants were asked to select the race of the applicant they had reviewed. Next, they were asked whether Michael had been to prison, his reason for conviction, and the number of years he had spent incarcerated. Participants were also asked whether a recommendation letter had been included in their application materials and, if so, who had written it.

**Demographics**

At the conclusion of our survey, we asked participants to respond to several questions regarding their demographic makeup. They reported their age, gender, race, and previous experience with hiring, and if they know anyone who had been incarcerated.

**Procedure**

Participants were instructed to read an informed consent agreement before beginning the experiment. By clicking the button to continue, they were consenting to the document. Participants were then randomly assigned to one of the sixteen conditions of applications. Participants read instructions saying, “You are an application analyst at a hiring agency in Indianapolis, Indiana. You have just received a new application and must rate the applicant’s qualifications. Please read over the application carefully and come to conclusions about the applicant based on the materials submitted by them.” Each application contained the same information except for the applicant’s criminal affiliation, a picture of their face, and whether they were presented with a recommendation letter. The information regarding skills, experience, and community involvement of the applicant remained consistent throughout each of the applications presented to the participants. There was also a cover letter submitted with each resume; the only difference was that the applicant who had been incarcerated disclosed the reason for his conviction.

Once they had completed looking over the application package they had been assigned, participants completed a brief questionnaire using 7-point Likert scales that measured how likely they were to recommend hiring this person, how they perceived the applicant’s competency and reliability, their perception of the applicant’s character, and their overall rating of the application. After the questionnaire,
participants answered questions about the name, gender, and criminal affiliation of the applicant, which were used as manipulation checks. Following this was a section of questions about the demographic data of the participant. Participants were compensated for their time via CloudResearch.

Results

Perceptions of Application Strength

We conducted a 4 (legal status: guilty vs. false confession vs. eyewitness vs. control) x 2 (race: White vs. Black) x 2 (letter of recommendation: present vs. absent) ANOVA to determine the effects of these variables on perceptions of application strength. We found a main effect for legal status, $F(3, 281) = 35.38, p < .001, \eta^2 = .27$. Participants viewed an application from a person who was guilty of committing the crime ($M = 3.83, SD = 1.91$) as weaker than an application from an eyewitness exoneree ($M = 5.80, SD = 1.24$), a false-confession exoneree ($M = 5.38, SD = 1.37$), or a member of the control group ($M = 6.00, SD = 1.10$). See Table 1 for all post hoc $t$-tests. See Figure 1 for graph.

Table 1. Post Hoc Test Results on Legal Status x Strength of Application ($N = 299$)

<table>
<thead>
<tr>
<th>Legal status</th>
<th>Legal status</th>
<th>$t$</th>
<th>$p$ (tukey)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>False confession</td>
<td>2.532</td>
<td>.057</td>
</tr>
<tr>
<td></td>
<td>Eyewitness</td>
<td>0.782</td>
<td>.863</td>
</tr>
<tr>
<td></td>
<td>Guilty</td>
<td>9.443</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>False confession</td>
<td>Eyewitness</td>
<td>-1.740</td>
<td>.305</td>
</tr>
<tr>
<td></td>
<td>Guilty</td>
<td>6.794</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Eyewitness</td>
<td>Guilty</td>
<td>8.587</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>
Perceptions of Competence

Recall that competence is a composite measure comprising measures of competency and the need for training. As mentioned, a composite of highly correlated (Cronbach’s $\alpha = .61$) question responses was used to evaluate this variable. We conducted a 4 (legal status: guilty vs. false confession vs. eyewitness vs. control) x 2 (race: White vs. Black) x 2 (letter of recommendation: present vs. absent) ANOVA on this composite score. We found a main effect for legal status, $F(3, 283) = 9.18, p < .001$, $\eta^2 = .09$.
\( \eta^2 = .09 \). Similar to the application-strength results, participants viewed guilty applicants \((M = 5.04, SD = 1.29)\) as less competent than eyewitness exonerees \((M = 5.81, SD = 0.01)\), false confession exonerees \((M = 5.59, SD = 1.03)\), and the control group \((M = 5.84, SD = 0.98)\). See Table 2 for all post hoc \(t\)-tests. See Figure 2 for graph.

Table 2. Post Hoc Test Results on Legal Status x Competency \((N = 299)\)

<table>
<thead>
<tr>
<th>Legal status</th>
<th>Legal status</th>
<th>( t )</th>
<th>( p ) (tukey)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>False confession</td>
<td>1.2574</td>
<td>.591</td>
</tr>
<tr>
<td></td>
<td>Eyewitness</td>
<td>0.0579</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Guilty</td>
<td>4.6791</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>False confession</td>
<td>Eyewitness</td>
<td>-1.1806</td>
<td>.640</td>
</tr>
<tr>
<td></td>
<td>Guilty</td>
<td>3.3829</td>
<td>.005</td>
</tr>
<tr>
<td>Eyewitness</td>
<td>Guilty</td>
<td>4.5540</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Figure 2. Competency x Legal Status

![Graph showing Competency x Legal Status](image)

_Note._ Error bars are standard errors.
The main effect of the letter of recommendation was significant, $F(1, 283) = 4.51, p = .04, \eta^2 = .02$. Applications that included a recommendation letter ($M = 5.69, SD = 1.07$) had a higher average rating of competency than those that did not have a letter of recommendation ($M = 5.47, SD = 1.11$). The main effect of race alone was not significant, $F(1, 283) = 0.53, p = .47, \eta^2 = .002$. The two-way interaction between the presence of a recommendation letter and legal status was not significant, $F(3, 283) = 0.57, p = .64, \eta^2 = .006$. The two-way interaction between the letter of recommendation and race was not significant, $F(1, 283) < 0.001, p = .996, \eta^2 < .001$. The two-way interaction between legal status and race was not significant, $F(3, 283) = 1.02, p = .38, \eta^2 = .01$. The three-way interaction letter of recommendation x legal status x race was not significant, $F(3, 283) = 0.39, p = .76, \eta^2 = .004$.

Perception of Character

Perception of character was measured as a composite score consisting of trustworthiness, the need for supervision, if the applicant would be welcomed, and participant’s comfort working with the applicant. We conducted a 4 (legal status: guilty vs. false confession vs. eyewitness vs. control) x 2 (race: White vs. Black) x 2 (letter of recommendation: present vs. absent) ANOVA to determine the effects of our independent variables on the participants’ judgments of the applicant’s character. We found a main effect for legal status such that participants viewed the guilty applicant ($M = 3.55, SD = 1.67$) as having a worse character than an eyewitness exoneree ($M = 5.69, SD = 1.09$), a false-confession exoneree ($M = 5.34, SD = 1.16$), and a member of the control group ($M = 6.08, SD = 0.79$), $F(3, 283) = 60.95, p < .001, \eta^2 = .392$. Interestingly, we also found that the perceived character of false-confession exonerees ($M = 5.34, SD = 1.16$) was significantly lower than perceptions of the control group ($M = 6.08, SD = 0.79$), $F(3, 283) = 60.95, p < .001, \eta^2 = .392$. See Table 3 for all post hoc $t$-tests. See Figure 3 for graph.
Table 3. Post Hoc Test Results on Legal Status x Character (N = 297)

<table>
<thead>
<tr>
<th>Legal status</th>
<th>Legal status</th>
<th>t</th>
<th>p (tukey)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control False confession</td>
<td>3.49</td>
<td>.003</td>
<td></td>
</tr>
<tr>
<td>Eyewitness Guilty</td>
<td>12.72</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>False confession Eyewitness</td>
<td>-1.55</td>
<td>.411</td>
<td></td>
</tr>
<tr>
<td>Guilty</td>
<td>9.13</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Eyewitness Guilty</td>
<td>10.71</td>
<td>&lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Character x Legal Status

Note. Error bars are standard errors.

The effect of the letter of recommendation alone was not significant, $F(1, 283) = 0.96, p = .33, \eta^2 = .003$. The effect of race alone was not significant, $F(1, 283) = 0.45, p = .50, \eta^2 = .002$. The interaction between the presence of a recommendation letter and legal status was not significant, $F(3, 283) = 0.54, p = .66, \eta^2 = .006$. The two-way interaction between the letter of recommendation and race was not significant, $F(1, 283) = 0.05, p = .83, \eta^2 < .001$. The two-way interaction between legal status and race
was not significant, $F(3, 283) = 2.20, p = .89, \eta^2 = .02$. The three-way interaction letter of recommendation x legal status x race was not significant, $F(3, 283) = 0.56, p = .64, \eta^2 = .006$.

Discussion

Many studies have focused on the discrimination of rightfully convicted individuals as they reintegrate into society (e.g., Bontrager et al., 2005; Greene & Dodge, 1995; Waldo, 2012), but many have neglected exonerees in this equation. The present study filled this gap by directly comparing employment discrimination and the perception of groups with varying degrees of criminal status. Contrary to our hypotheses, the applicant’s race did not have a significant effect on the evaluation of their application package. A few explanations are possible here. The first is the fact that a photograph of the applicant is not typically submitted with application materials and thus the participants in our study may have realized that race was an important manipulation in the experiment, which could skew the results that we collected. Second, there is some previous research showing that presenting a picture of an individual could humanize them and lead to more positive interactions by reducing anonymity (Teubner et al., 2014). This is possible in the case of this study, as there were high ratings shown, on average, of most applicants. Though a nonsignificant effect of race is an important one to discuss based on the results, the most notable difference came in the perceived character of the applicant.

The applicant who was truly guilty of committing a crime was consistently judged more negatively than those who were presumed innocent, contrary to our hypothesis that exonerated individuals would be rated similarly to guilty individuals. This finding replicates many previous studies that have found this discrimination against people who had committed a crime in the past; however, it is just as important to study the discrimination faced by people who have been wrongfully imprisoned. There has been an increase in studies on wrongful convictions, though there has been much less research on the varying contributing factors that led to those wrongful convictions.

In agreement with our hypotheses, the present study found that the reason for conviction does have an impact on the public’s perception of exonerees. Participants reported believing that those who falsely confessed had weaker character than those who had not been incarcerated. There was no such difference found between those who were imprisoned based on faulty eyewitness testimony and the control group. From this, it may be the case that the public is not educated enough on the causes of false confessions. There is sufficient research showing that situational factors such as interrogation tactics can elicit false confessions from suspects (Kassin et al., 2010). Increased public awareness and education about these tactics can create a universal
understanding of these instances and alleviate the negative perception of people who falsely confess to crimes for a variety of reasons.

The findings of this study highlight the prejudice that continues to exist against people who have been incarcerated. Though there has been a recent push to allow previously incarcerated individuals more second chances, there is still a long way to go before these practices are implemented consistently. Legislative progress is important in this field, though bias from those in the workforce must also be reduced to create a safe and equitable environment for formerly incarcerated individuals.

**Policy and Practice Implications**

The current legislation around reentry services for exonerees is inconsistent and time-costly. Specifically, most exonerees are released from incarceration with no housing, transportation, or health insurance and often still have the convictions on their records (Witness to Innocence, 2019). As of June 2022, 38 states and the District of Columbia had statutes in place to compensate exonerees for the time they were incarcerated, though fewer than half of the exonerees who sought this compensation were awarded any money at all (National Registry of Exonerations, 2022). In some states, reentry services for exonerees are not afforded when individuals are thought to have contributed to their convictions, namely by pleading guilty or falsely confessing (National Registry of Exonerations, 2022). These obstacles, and more, await wrongfully convicted individuals after they are exonerated of their crimes. It is important that states assure reentry services for exonerated people, as exonerees face similar challenges to those who were rightfully convicted and then released.

Nearly one-third of exonerated individuals do not have their records expunged, which is a predictor of these individuals committing crimes after they are released (Shlosberg et al., 2014). As shown in the present study, the presence of a criminal record can negatively affect the employment chances of an applicant. A streamlined process should thus exist for exonerated individuals to have their records expunged promptly. Although some states have straightforward processes for exonerees to clear their records, in many more, it can take years for an exonerated person’s record to be expunged (Innocence Project, 2023). Creation of a consistent and streamlined process to remove wrongful charges from exonerated individuals must become a federal priority. A policy change of this magnitude can create more opportunities for exonerees after their release and can reduce the number of exonerees who do commit crimes after they have reentered society.

Moreover, rightfully convicted individuals face discrimination after serving their sentences. It is important to implement proactive policies at a federal level that afford formerly incarcerated individuals the same opportunities as other citizens. These policies could include creation of programs in prisons that ready incarcerated
individuals for employment upon release. Instead of subjecting people who are incarcerated to an isolated existence, cultivating marketable skills for them to use after release can positively influence their reentry journeys (Esperian, 2010).

Limitations and Future Research

The most prominent limitation of this study is the participant sample. Just 139 of the 299 (46.49%) participants who completed the survey reported that they have been in charge of hiring decisions at their place of employment. The full set of data may therefore not be representative of actual employers and may be more representative of the population in general. The racial makeup of the participants was also not very diverse (83.3% White), which may have had an effect on our results. Moreover, though an authentic recreation of an application process was attempted, it is not as genuine as a real application process. It may increase the external validity of the results to submit applications to active job postings and analyze the rate of response from potential employers, as previous research has done with housing applications (Kukucka et al., 2021).

One potential limitation is the perceived age of the man in the picture shown with the applications. In the pilot study, participants were shown the images of two men, one Black man and one White man. There was a significant difference found in attractiveness \((p < .001)\), so a second pilot study was run with two new pictures of one Black man and one White man. The second pilot study found a significant difference in intimidation \((p = .001)\). In comparing the two results together, the Black man in the first study and the White man in the second study were not significantly different in attractiveness \((p = .33)\) or intimidation \((p = .42)\); however, there was a significant difference in the age of the man pictured \((p = .005)\), though it was determined that because the variable was categorical and the majority of respondents indicated that the men were in the same age range of 26 to 35 years old (62.7% and 74.5%), this result may not have been significant enough to skew the results of the study.

Additionally, submitting images with the application may not be the most effective way to vary the race of the applicant. There was no effect of race found in our study, which is a stark contrast to previous research findings. This could be because participants were able to infer that race was being varied in the application materials, because a picture is not typically submitted with an application. It may be the case that submitting an image may have humanized all applicants so they received higher overall ratings than they may have if the pictures had not been included. To address this possible issue, future studies can vary the race of the applicant by changing the name to have a traditionally White name and a traditionally Black name for the applicant. The subtle change of name has been shown to be effective in varying the race of applicants in past research (Varghese et al., 2010) and could be useful in exoneration research.
Future research may include studying the differences in application evaluations between male and female exonerees. There has been some research done on the gender differences of incarcerated individuals and their treatment in prison (Carcedo et al., 2008; Toman, 2017), though there is little on exonerees, specifically. As suggested before, sending applications with varying legal statuses to actively open positions may create a more generalizable result from future studies. Obtaining further information about employers in future studies may clarify why they make the decisions and evaluations that they do. This includes diversifying the sample of participants used in studies in this field.

Conclusion

Overall, the present study has made clear the bias that still exists toward previously incarcerated people, including those who were wrongfully convicted, specifically those who falsely confessed. Formerly incarcerated individuals face many obstacles when they are released from prison. These obstacles may stem largely from affiliation with the criminal justice system, not just from the fact that a crime was committed. As many organizations have begun pointing out the flaws in our criminal legal system and setting innocent people free, it is clear that a great need exists for policy and program improvement upon exonerees’ release. It is important to educate the public so society can set aside their prejudice and see the human first, instead of the conviction.
References


All of Us or None. (n.d.). *About: The Ban the Box campaign*. Ban the Box. Retrieved February 1, 2023, from http://bantheboxcampaign.org/about/#.Y9sE1OzMI-Q


Abstract

Work-family life balance influences employee job satisfaction and quality of life. When work and family life are in balance, employees perform at higher levels and have fewer health problems. U.S. Army Reservists in both civilian and Reserve medical occupations may face work-life balance conflicts from both their full-time/civilian and part-time/military medical occupations. Army Reservists in a medical unit were surveyed about their work-family life balance behaviors using the Work–Life Climate Scale in an exploratory cross-sectional study. Most reported rarely experiencing work-life conflict behaviors; however, many also reported that they worked late three to seven days of the week. Long work hours decrease work-family balance levels and may lead to lower quality of life and poor health outcomes. If results of larger, more-definitive studies also note long work hours as a work-family life conflict for Reservists in medical units, organizational and policy change may be needed to support the mission.

Introduction

Work-life balance—high life satisfaction and low conflict across work and nonwork roles—is affected by the interaction of individual and workplace elements. On one hand, when role conflict arises, worker stress increases and job satisfaction decreases, leading to poor mental and physical health outcomes. On the other hand, positive participation in one role can spill over to the other role. Balancing work and nonwork roles effectively with high participation in and integration of both roles can lead to improved quality of life (Knecht & Freund, 2016; Rahman et al., 2017; Sirgy & Lee, 2018).

Specifically, the interaction between work and family life can contribute positively or negatively to both job and family satisfaction, which are influencers of overall quality of life and health. A review found high employee job and life performance and satisfaction levels as well as low levels of physical and mental health
problems when work and home life were balanced with active participation in both roles (Sirgy & Lee, 2018). Incompatibility of work and home life because of long, inflexible, or shift-work hours, job demands, or work stress was shown to increase work-family conflict levels and decrease family life satisfaction (Ford et al., 2018). Life becomes more unbalanced in work settings that possess high job expectations, lack of time to meet those expectations, low autonomy and schedule flexibility, role uncertainty (Sirgy & Lee, 2018), and job expectation of being available to work during nonwork hours (Dettmers, 2017).

Finding the balance in managing dual roles of work and family has been suggested as a strategy that can enrich both roles (Leduc et al., 2016). Work skills can enrich family skills, and vice versa. It seems that preference for boundary management by segmenting boundaries, clearly separating work life from family life, may not always prevent conflict. Integrating or interfacing boundaries predicted positive work-family life enrichment but was also associated with more conflict. It has been speculated that more work-family conflict is experienced when work disrupts family life (Leduc et al., 2016).

For the workplace, there are positive support and policy influences on work–family life balance, including flextime, childcare assistance, social support, and worksite health-promotion programs, leading to improved recruitment, retention, and productivity (Sirgy & Lee, 2018). For the employee, positive workplace supports and work–family life-promoting policies show small effects on work-family conflict; however, employee perception of workplace efforts supportive of work–family life balance were found to be a predictor of job and family life satisfaction levels (Ford et al., 2018).

Work–Family Life Balance in the Military

Challenges for the military around work-family life balance are different from civilian issues in their intensity and demands, as the military is a 24/7 operation with heavy workloads and accompanying work-family conflict stressors. For example, relocations and separations with little advance notice are common. Exposure to combat, trauma, and other hazards during deployment are related not only to mental and physical health problems for the soldier but also to family stress and conflict in accommodating for the absence (Wadsworth & Southwell, 2011). Additionally, in a literature review of nondeployment contributors to soldier health outcomes, work–family life balance and role conflict were noted as health risks (Brooks & Greenberg, 2018). In response, the military has adopted more flexible and family-friendly benefits over time to support the mission (Wadsworth & Southwell, 2011).

In the military, work-family life balance requires serving the Army’s demands for time and commitment while still serving self and family. When active-duty soldiers
were surveyed regarding their and their families’ needs, work-life balance was a priority concern noted, especially for officers and those stationed away from family. Those who identified this as a top priority requested that recreation and family bonding activities be provided by the military (Sims et al., 2017).

Medical and healthcare professionals also face occupational stressors and challenges to work-life balance related to work setting and organizational factors that may lead to burnout, physical and mental exhaustion. In healthcare, work-life balance is also challenging, and employee coping behaviors have not been frequently measured. The conflicting demand of sacrificing personal needs for patient and safety needs was shown to lead to occupational stress and poor family life satisfaction. In a large study of healthcare professionals in a health system, low levels of work-life balance were related to occupational stress and burnout (Schwartz et al., 2019).

The U.S. Army Reserve, consisting of civilian soldiers with careers or attending college, fills the positions of active-duty soldiers while they are deployed overseas, responding to natural disasters, safeguarding communication technology, providing biohazard training, and working in civil affairs and military and medical services support (U.S. Army Reserve, n.d.). Those Reservists serving in both civilian and Reserve medical occupations may face work-life balance conflicts.

The research question in this exploratory study asks, “What is the work-life balance level of soldiers in a medical unit of the U.S. Army Reserve stationed in a Midwestern state?”

Methods

Sample

Twenty-nine U.S. Army Reservists from a medical unit in a Midwestern state were asked to participate in a survey research study before a scheduled battle assembly (monthly trainings in which Reservists practice military readiness skills in case of eventual mobilization). Their Reserve roles included nurses, veterinarians, emergency medical technicians, and other allied health professions. All Reservists in attendance at a 2023 battle assembly were asked to participate in the study and were provided an informed consent document detailing the information about which they would be surveyed and their freedom to decide whether to participate. All consented. Most (59%) were female, and a little over half (55%) were never deployed. See Table 1 for participant demographic information.
Table 1. Demographics of the Sample

<table>
<thead>
<tr>
<th>Demographic</th>
<th>n(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>11(37.9)</td>
</tr>
<tr>
<td>Female</td>
<td>17(58.6)</td>
</tr>
<tr>
<td><strong>Deployments</strong></td>
<td></td>
</tr>
<tr>
<td>No previous deployments</td>
<td>16(55.2)</td>
</tr>
<tr>
<td>One deployment</td>
<td>5(17.2)</td>
</tr>
<tr>
<td>Two or more deployments</td>
<td>7(24.1)</td>
</tr>
</tbody>
</table>

Note. N = 29. One participant was removed because it appeared there was purposeful improper reporting. *Totals not equaling 100% indicate missing data.

Instrument

The Work–Life Climate Scale (Sexton et al., 2017), with good internal consistency (Chronbach alpha = .83) and strong psychometrics, was used to measure work-family life balance behaviors in respondents. Following a prompt that asked about past week frequency of occurrence, seven work-family life balance behavioral statements such as “Skipped a meal” and “Changed personal plans because of work” were listed. Table 2 shows all questions. Respondents were to rate on a scale of rarely (<1 day), some (1–2 days), occasionally (3–4 days), all (5–7 days), or not applicable. Two demographic questions, asking about sex and number of deployments, were also asked.
Table 2. Descriptive Statistics for Individual Items on the Work–Life Climate Scale

<table>
<thead>
<tr>
<th>During the past week, how often did this occur?</th>
<th>Rarely (&lt;1 day) n(%)</th>
<th>Some (1–2 days) n(%)</th>
<th>Occasionally (3–4 days) n(%)</th>
<th>All (5–7 days) n(%)</th>
<th>Not applicable n(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skipped a meal</td>
<td>11(37.9)</td>
<td>9(31.0)</td>
<td>2(6.9)</td>
<td>6(20.7)</td>
<td>1(3.4)</td>
</tr>
<tr>
<td>Ate a poorly balanced meal</td>
<td>10(34.5)</td>
<td>11(37.9)</td>
<td>3(10.3)</td>
<td>5(17.2)</td>
<td>—</td>
</tr>
<tr>
<td>Worked through a shift without a break</td>
<td>8(27.6)</td>
<td>9(31.0)</td>
<td>5(17.2)</td>
<td>4(13.8)</td>
<td>3(10.3)</td>
</tr>
<tr>
<td>Arrived home late from work</td>
<td>7(24.1)</td>
<td>7(24.1)</td>
<td>10(34.5)</td>
<td>3(10.3)</td>
<td>2(6.9)</td>
</tr>
<tr>
<td>Slept less than 5 hours/night</td>
<td>12(41.4)</td>
<td>7(24.1)</td>
<td>6(20.7)</td>
<td>4(13.8)</td>
<td>—</td>
</tr>
<tr>
<td>Changed personal/family plans because of work</td>
<td>13(44.8)</td>
<td>8(27.6)</td>
<td>3(10.3)</td>
<td>3(10.3)</td>
<td>2(6.9)</td>
</tr>
<tr>
<td>Felt frustrated by technology</td>
<td>11(37.9)</td>
<td>6(20.7)</td>
<td>5(17.2)</td>
<td>5(17.2)</td>
<td>2(6.9)</td>
</tr>
</tbody>
</table>

Procedure

In a cross-sectional study, after all approvals and consent, and immediately before a weekend battle assembly, respondents completed the confidential written survey.

Analysis

Frequencies and proportions were calculated on each of the items.

Results

As seen in Table 2, for five of the seven items, most respondents reported frequency of occurrence as rare; however, for the item “ate a poorly balanced meal,” most responded that this occurred some days. About 28% of respondents, though, “skipped a meal” or “ate a poorly balanced meal” either occasionally or all days of the week.

For the item “arrived home late from work,” most reported that this occurred occasionally. Almost 45% of respondents, though, reported that they “arrived home late from work” either occasionally or all days of the week.

About 35% also reported that they “slept less than five hours per night” or “felt frustrated by technology” either occasionally or all days of the week.

Discussion

Respondents were medical and healthcare professionals working in both civilian and military roles. Their roles in healthcare demand serving patient needs, which may lead to poor work-life balance (Schwartz et al., 2019), and their roles in the military demand serving the Army’s needs, which may take time away from family (Sims et al., 2017). Because of the dual demands, Reservists in medical units may demonstrate poor work-family life balance behaviors. Results of this study, however, indicate few perceived work-family life conflicts of respondents. Possibly, these respondents positively integrated their work and life roles, as both were healthcare roles that served others. Balancing their roles in this way may lead to improved life satisfaction and quality of life (Knecht & Freund, 2016; Rahman et al., 2017; Sirgy & Lee, 2018).

In addition, many respondents had had one or no deployments. Exposure to combat increases soldier stress and decreases work-family life balance as the family accommodates for the soldier’s absence (Wadsworth & Southwell, 2011). Being Reservists, not full-time soldiers, may also have contributed to their reporting of few poor work-family life balance behaviors. Studies on active-duty soldiers reported
poor work-life balance as a health risk (Brooks & Greenberg, 2018) and a priority concern for soldiers and their families (Sims et al., 2017).

Almost half of respondents, however, reported that they “arrived home late from work” either occasionally or all days of the week. Long work hours may be a potential work-life balance and health concern for this group. Especially for those in healthcare, poor work-life balance was associated with job burnout and low family-life satisfaction (Schwartz et al., 2019). Of all the work-family life behaviors, those that disrupt family time cause the most conflict (Leduc et al., 2016). Work and life become most unbalanced when the employee is expected to work during nonwork hours (Dettmers, 2017).

Although the survey gathered perceptions of an understudied group as a precursor to a possible larger study, small sample size, geographic location, and possible social desirability bias may have affected generalizability of results. If results of larger, more-definitive studies, including qualitative examination, point to long work hours as a work-family life conflict for Reservists in medical units, more organizational efforts and policy change may be needed to support the mission; therefore, continuation of the military’s family-friendly policies (Wadsworth & Southwell, 2011) may be appropriate. Other organizational changes at the workplace, such as flextime and childcare programs, positively perceived by employees in business and industry and showing some effects (Ford et al., 2018), should also be investigated. When work and family life are in balance, employees perform at higher levels, are more satisfied with their jobs, and have fewer health problems (Sirgy & Lee, 2018).

Additionally, as work-life balance is often a multivariate, complex topic, perhaps a more-qualitative approach could be used in the future in this population to identify aspects of work-life balance not captured in this questionnaire approach. Through facilitation of a qualitative, more open-ended discussion and data-collection process, perhaps novel work-life balance initiatives can be initiated to meet needs that remain undiscovered within medical units of U.S. Army Reservists today.
References


EFFECTS OF DISSOLVED OXYGEN ON NORTHERN LEOPARD FROG (LITHOBATES PIPiens) GASTROCNEMIUS MUSCLE RECRUITMENT AND FATIGUE

EVA L. GATES, PENN STATE ALTOONA
JOSEPH P. LEVRI, PENN STATE ALTOONA
MENTOR: LARA LADAGE

Abstract

Dissolved oxygen levels in aquatic systems have decreased because of temperature increases caused by climate change, which in turn has affected ecosystems and wildlife. Many physiological processes in aquatic organisms require a certain dissolved-oxygen range, and decreasing levels can compromise proper functioning. Previous studies have linked muscle performance to dissolved oxygen levels in a variety of aquatic species, but less research has been dedicated to amphibians. Because many amphibians engage in cutaneous respiration, especially when dwelling in aquatic habitats, dissolved oxygen levels may have a significant impact on muscle performance in this taxon. This experiment investigated the effects of dissolved oxygen and time in vitro on frog skeletal muscle contractile force and fatigue. Results did not vary significantly when dissolved oxygen was altered, but fatigue and contractile force experiments did correlate with time in vitro. Although we did not find an effect of dissolved oxygen levels on muscle characteristics in vitro, a better understanding of the effects of dissolved oxygen on muscle performance, particularly in vivo, could be beneficial as climate change alters the oxygen content of aquatic systems, with the potential to affect physiology and behavior.

Keywords: cutaneous respiration, dissolved oxygen, contractile force, fatigue, in vitro

Introduction

It is known that the solubility of oxygen is temperature-dependent and that bodies of water with lower temperatures will have a higher dissolved oxygen content than those of higher temperatures. Both fresh- and saltwater have recently been observed to fluctuate in dissolved oxygen, and researchers have attributed this to climate change and nutrient over-enrichment (Irby et al., 2018; Whitehead et al., 2009). Specifically, the proportional increase in stream temperature with air temperature as a result of climate change has caused decreases in dissolved oxygen.
levels (Ficklin et al., 2013). Similarly, nutrient runoff from agriculture in the Mississippi basin affects downward bodies of water and ecosystems, reflected by decreased dissolved oxygen in the Gulf of Mexico (Joyce, 2000). The decreased dissolved oxygen levels caused by increased stream and air temperatures as a result of climate change can have severe effects on biodiversity, behavior, and distribution of wildlife inhabiting these bodies of water. This includes decreased growth rates of wildlife in hypoxic waters and, in some extreme hypoxic conditions, creation of aquatic dead zones that cannot sustain life (Breitburg, 2002; Joyce, 2000; Waldrop et al., 2018). With increasing concerns of the effects of increased air and water temperatures and the decrease in dissolved oxygen content in these waters on aquatic ecosystems and wildlife, further investigation of the relationship between oxygen availability and organism performance can provide insight on the consequences of decreased dissolved oxygen levels.

Dissolved oxygen can contribute to population-level variation in aquatic systems, such as reduced growth and development (Breitburg, 2002), which, mechanistically, may in part be due to the relationship between oxygen availability and the underlying physiology of an organism. Oxygen transport capacity has effects on metabolic rate, as hypoxic conditions are negatively correlated with metabolic activity in some reptiles and amphibians (Gangloff & Telemeco, 2018). This is particularly evident in the relationship between oxygen availability and muscle performance and fatigue. Muscle performance is sustained through a consistent supply of oxygen for oxidative phosphorylation and adenosine triphosphate (ATP) genesis (Wittenberg and Wittenberg, 1989). Muscle fatigue can be attributed to ATP availability, in that if muscle activity exceeds ATP synthesis rates, then fatigue is observed (Sundberg & Fitts, 2019). In hypoxic conditions, muscle fatigue has been shown to occur earlier (Amann et al., 2006; Verges et al., 2010), as decreased oxygen is directly related to ATP depletion (Sahlin et al., 1998). The effects of dissolved oxygen levels on the muscle physiology of aquatic animals are well documented (Waldrop et al., 2018). For example, aquatic species such as whiteleg shrimp (Litopenaeus vannamei) decrease distance traveled per tail-flip, and crucian carp (Carassius carassius) have shown a decrease in maintained fastest swimming speed in low-oxygen environments (Domenici et al., 2013; Duan et al., 2022; Penghan et al., 2014). Overall, these studies demonstrate the importance of dissolved oxygen levels on muscle-based performance in aquatic environments.

Much of the research on dissolved oxygen in aquatic systems and muscle performance has been conducted in fish and invertebrates, with comparatively less research devoted to the effect of dissolved oxygen concentration on muscle performance in amphibians. This is significant because many amphibians engage in some degree of cutaneous respiration, using their skin as a membrane for gas exchange. While lungs and gills (in amphibian larvae) generally provide both a thinner surface and increased surface area for gas exchange, making it preferable for gas
exchange, many amphibians can take in as much as 30% of their oxygen through the skin, particularly when metabolic rate is lower (Feder & Burggren, 1985). Additionally, some amphibian species have adaptations to increase the surface area of the skin, while other species increase capillary perfusion during high respiratory activity, such as mating rituals (Feder & Burggren, 1985). This suggests that cutaneous respiration has a functional role in oxygen uptake and may be affected by levels of dissolved oxygen, potentially affecting muscle performance as well; however, this concept has not been well studied.

Because oxygen is required for proper muscle performance, and declining dissolved oxygen may modulate this relationship, we investigated the effects of dissolved oxygen on in vitro skeletal muscle contractile force and fatigue in northern leopard frogs (*Lithobates pipiens*). We hypothesized that bathing a muscle with varying dissolved oxygen levels would affect muscle performance properties. Specifically, we predicted that decreased dissolved oxygen would decrease contractile force and decrease the time to fatigue.

**Materials and Methods**

Fifteen northern leopard frogs were acquired from Carolina Biological Supply (Burlington, North Carolina). This species is known to respire cutaneously when hibernating in water (Tattersall & Boutilier, 1997). Frogs were housed in pairs in a plastic enclosure (34.5 cm x 30 cm x 13 cm) containing rocks and water. The frogs were each fed 8–10 mealworms a day and were held in captivity for up to two days. All procedures were approved by the Pennsylvania State University’s Institution Animal Care and Use Committee (protocol #201546456).

The frogs were humanely sacrificed, and the gastrocnemius muscle was separated from the tibiofibular bone at the insertion point. Ringer’s solution (200–250 mOsm, 6.6 g/L NaCl, 0.15 g/L KCl, 0.15 g/L CaCl$_2$, 0.2 g/L NaHCO$_3$) was used to keep the muscle moist after removal. The femur was left attached to the tibiofemoral joint, keeping it connected to the gastrocnemius at the muscle origin. A thread was tied around the interface of the muscle and Achilles tendon, with the other end of the string attached to a force transducer to record muscle force. To apply a stimulus voltage, a stimulating rod was attached to a ring stand and positioned so it was in contact with the middle of the muscle.

Both legs from each frog were used in the experiment, one as treatment and one as control. We systematically alternated which leg muscle (right or left) went into the control and treatment group for each frog to ensure that any differences in muscle performance were not due to whether the left or right leg was used. For the control group, gastrocnemius muscles were continuously bathed in oxygenated Ringer’s solution at room temperature (21 ± 0.5 °C). The solution was oxygenated with a
bubbler and had an average dissolved oxygen reading of 8.0 ± 0.2 mg/L, which represents a normoxic environment for this species (Tattersall & Boutilier, 1997). For the treatment group, gastrocnemius muscles were continuously bathed in standard Ringer’s solution at room temperature (21 ± 0.5 °C). This solution was not bubbled and had an average dissolved oxygen reading of 4.5 ± 0.2 mg/L, which represents hypoxic conditions (Tattersall & Boutilier, 1997).

A PowerLab data acquisition system (AdInstruments, Colorado Springs, CO) was used to control the stimulator rod and analyze the data collected through the force transducer. Contractile force was measured in Newtons, and the preload of each gastrocnemius muscle tested was set at 0.75 N, to maintain preload consistency across all muscles. The effects of voltage stimulus amplitude and duration of stimulation (fatigue) on contractile force were observed in relation to dissolved oxygen treatment. The effect of voltage stimulus amplitude on contractile force was investigated by using a macro program that stimulated the muscle once per second, in increments of 0.05 V, from 0.05 to 1 V, for 20 seconds. The voltage required to elicit the first detectable contractile force was recorded, as well as the strength of the first contraction. The final contractile force after stimulation with 1 V was also measured. We used a different macro program to then induce muscle fatigue. Fatigue was measured by first assessing the voltage that induced maximal contractile force, then continually stimulating the muscle at that voltage for 30 seconds. The measured force at the end of the 30 seconds was recorded, as well as the time required for the contractile force to reach zero, if it reached zero before the end of 30 seconds. The maximum contractile force reached across the 30-second interval and the time at which the muscle reached maximum contractile force were also recorded. Both the effect of voltage stimulus amplitude and fatigue variables were measured on the same day, at 11 AM, 3 PM, and 7 PM, to determine if the dependent variables were affected by time in vitro.

We used a repeated measures analysis of variance (RM ANOVA) to analyze the effects of dissolved oxygen treatment and time point on all dependent variables. All analyses were conducted with SPSS for Windows, v. 27 (IBM Corp.), and we considered results to be statistically significant if p ≤ 0.05.

Results

Data for the 3 PM time point were transformed but could not conform to the assumption of homogeneity of variance (p < .05), so the data at this time point were not used in the final parametric analyses. Box’s test of equality of covariance matrices indicated conformation to equality of covariances for most variables (all p > .055), but the force of first detectable contraction and the force at 1 V had to be log-transformed to conform to the assumption of homogeneity of variances (after transformation, p = .148 and .202, respectively).
Stimulus Voltage to Elicit First Detectable Contractile Force

We found that voltage needed to elicit the first detectable contractile force increased between 11 AM and 7 PM \( (F_{1,26} = 5.956, p = .022) \) but there were no statistically significant effects of dissolved oxygen treatment \( (F_{1,26} = 0.019, p = .184) \) or the interaction between time point and dissolved oxygen treatment \( (F_{1,26} = 0.756, p = .393) \) (Figure 1).

Figure 1. Voltage Required to Elicit First Detectable Contractile Force

Note. No significant difference existed between the control (normoxic) and hypoxic oxygen treatments \( (p = .184) \) or in the interaction between dissolved oxygen treatment and time \( (p = .393) \); however, the voltage to elicit the first detectable contractile force differed between tests done at 11 AM and those done at 7 PM \( (p = .022) \).

Force of First Detectable Contractile Force

We found that the force of the first detectable contractile force decreased between 11 AM and 7 PM \( (F_{1,26} = 7.532, p = .011) \) but there were no statistically significant effects of dissolved oxygen treatment \( (F_{1,26} = 0.192, p = .665) \) or the...
interaction between time point and dissolved oxygen treatment ($F_{1,26} = 0.263, p = .613$) (Figure 2).

**Figure 2. Force of First Detectable Muscle Contraction**

![Chart showing force of first detectable muscle contraction at 11 AM and 7 PM with error bars.

Note. No significant difference existed between the control (normoxic) and hypoxic oxygen treatments ($p = .665$) or in the interaction between dissolved oxygen treatment and time ($p = .613$); however, the force of the first detectable contraction differed between tests done at 11 AM and those done at 7 PM ($p = .011$).

**Force at 1 V Stimulation**

We found that the force at 1 V stimulation decreased between 11 AM and 7 PM ($F_{1,27} = 4.634, p = .040$) but there were no statistically significant effects of dissolved oxygen treatment ($F_{1,27} = 0.273, p = .606$) or the interaction between time point and dissolved oxygen treatment ($F_{1,27} = 1.152, p = .293$) (Figure 3).
Figure 3. Contractile Force at Stimulation of 1 V

Note. No significant difference existed between the control (normoxic) and hypoxic oxygen treatments ($p = .606$) or in the interaction between dissolved oxygen treatment and time ($p = .293$); however, the force at stimulation of 1 V differed between tests done at 11 AM and those done at 7 PM ($p = .040$).

Muscle Fatigue: Contractile Force at the End of 30 Seconds

We found that contractile force at the end of 30 seconds of stimulation decreased between 11 AM and 7 PM ($F_{1,18} = 5.950, p = .025$) but there were no statistically significant effects of dissolved oxygen treatment ($F_{1,18} = 0.181, p = .676$) or the interaction between time point and dissolved oxygen treatment ($F_{1,18} = 0.120, p = .733$) (Figure 4).
Figure 4. Contractile Force at 30 Seconds of Stimulation (Fatigue)

Note. No significant difference existed between the control (normoxic) and hypoxic oxygen treatments ($p = .676$) or in the interaction between dissolved oxygen treatment and time ($p = .733$); however, the time to reach force at 30 seconds of stimulation differed between tests done at 11 AM and 7 PM ($p = .025$).

Muscle Fatigue: Time Until Fatigue

We found that time until fatigue decreased between 11 AM and 7 PM ($F_{1,18} = 6.111$, $p = .024$) but there were no statistically significant effects of dissolved oxygen treatment ($F_{1,18} = 0.062$, $p = .806$) or the interaction between time point and dissolved oxygen treatment ($F_{1,18} = 0.760$, $p = .395$) (Figure 5).
Figure 5. Time for Muscle to Reach Full Fatigue

![Bar chart showing time for muscle to reach full fatigue at 11 AM and 7 PM]

*Note.* No significant difference existed between the control (normoxic) and hypoxic oxygen treatments ($p = .806$) or in the interaction between dissolved oxygen treatment and time ($p = .396$); however, the time to reach full fatigue differed between tests done at 11 AM and those done at 7 PM ($p = .024$).

**Discussion**

For contractile force with increasing stimulation, the voltage required to elicit the first detectable gastrocnemius contraction, strength of the first detectable contraction, and contractile force at 1 V did not show significant changes based on dissolved oxygen treatment. These values changed over time, however; after 8 hours, a more intense stimulus was required to elicit the first contractile force, and there was a decrease in the first detectable contractile force and force at 1 V stimulation. For variables related to fatigue, the amount of time required for gastrocnemius muscles to fatigue and the contractile force after 30 seconds of stimulation did not differ between the control and low dissolved-oxygen treatment groups; however, contractile force after 30 seconds of stimulation and time until fatigue did significantly decrease after 8 hours in vitro. While our hypothesis of decreased muscle performance due to a decrease in dissolved oxygen was not supported, there was a significant difference between tests done later in the day, which resulted in a decline of muscle performance after 8 hours in vitro.

Decreased strength of contraction and increased fatigue over time are consistent with previous studies showing that muscles in vitro are viable for...
producing adequate contractile force for only a few hours (Smith & Meyer, 2020). Muscle viability loss is likely due to the lack of blood flow and vasculature that supply muscles with oxygen and nutrients necessary to carry out metabolic processes while also removing waste products (Smith & Meyer, 2020). Viability is not the only factor, as studies have shown that contractile force was also reduced significantly after 40 minutes of incubation of in vitro muscles (Croes & von Bartheld, 2007). Decreases in muscle viability and the strength of contractions over time have shown to be less severe in muscles incubated in oxygenated Ringer’s solutions (Croes & von Bartheld, 2007), but further research concerning the viability of muscles over time for measuring contractile force and fatigue may be beneficial. It is likely that ischemic conditions are the cause of decreased muscle performance and viability over time and could possibly be prevented by performing this study in vivo with a continuous blood supply.

We observed no effect of the use of a hypoxic Ringer’s solution in decreasing muscle viability, nor in decreasing the strength of muscle contractions or increasing fatigue. The ineffectiveness of the dissolved oxygen treatment on the variables measured in this study is consistent with some other studies that used similar techniques. For example, Croes and von Bartheld (2007) did not report significant effects of the use of oxygenated Krebs buffer on twitch force in lateral gastrocnemius and superior oblique muscles of juvenile white leghorn chickens (Croes and von Bartheld, 2007). In comparison, some species have exhibited an increase in muscle performance with the use of oxygenated systems. Atlantic cod (Gadus morhua), for example, exhibited a 41% decrease in swimming speed under deep hypoxic conditions (4.3 kPa of oxygen; Herbert & Steffensen, 2005), which could be related to a decrease in the strength of muscle contractions. Similarly, golden grey mullet (Liza aurata) demonstrated a 25% decrease in escape behavior in response to predatory stimuli under hypoxic conditions (10% air saturation; Lefrançois et al., 2005), which could be attributed to a decrease in muscle metabolic activity. In the previous two studies, the difference between normoxic and hypoxic was much greater than in our study, which may explain why they arrived at significant differences whereas we did not. It may be that we did not have sufficient resolution to distinguish between our two dissolved-oxygen treatment groups.

The inconsistency of our data with other studies could be attributed to other differences in experimental methodology. For example, in our study, muscles were wrapped in paper towels soaked with normoxic or hypoxic Ringer’s solution and bathed during testing. In other studies, the experiment was performed while the muscles were submerged in the Ringer’s solution (Croes & von Bartheld, 2007), which may have provided a more consistent dissolved-oxygen environment to the muscles. There were also differences in the methodology for oxygenating or deoxygenating the Ringer’s solution; there has not been a universally accepted method for manipulating oxygen in this type of experiment. Croes and von Bartheld (2007) used a 95% O₂ and
5% CO₂ gas to oxygenate the buffer, whereas we used atmospheric gas, which is about 21% O₂. This compositional difference could have affected oxygen or other dissolved gas levels of Ringer’s solutions, thus affecting muscle performance. Another technique to create a hypoxic environment is to bubble the Ringer’s solution with nitrogen and seal the reservoir where the muscle is housed (Penghan et al., 2014). The method of deoxygenating rather than oxygenated solutions is another source of variation that may have contributed to the differences in the effect of hypoxia found in our and other studies.

It is possible that the metabolic rate of the frog is a major contributor to the extent of cutaneous respiration. The proportion of cutaneous oxygen uptake may positively correlate with metabolic rate. It is also plausible that frogs may rely on cutaneous respiration only when submerged for a longer time or when they are exposed to warmer temperatures that limit the dissolved oxygen content of water. Instead, the lungs may be the primary source of respiration when the amphibian is warm or breathing air. If any of the previous are important modulators among dissolved oxygen, cutaneous respiration, and muscle performance, simply exposing a disembodied muscle to dissolved oxygen at room temperature may not change the characteristics of the muscle.

Decreased dissolved oxygen in aquatic ecosystems due to climate change has been affecting wildlife distribution and behavior. Specifically, variation in dissolved oxygen alters the distribution of aquatic populations, leading to behavioral and physiological changes in some species and disrupted food webs on a larger scale (Breitburg, 2002). Additionally, nitrogen concentrations in some aquatic systems are fluctuating because of human addition of fixed nitrogen from agriculture (Joyce, 2000), which in turn affects oxygen levels. Climatic and anthropogenic environmental change are affecting more and more aquatic ecosystems, and insight on how dissolved oxygen levels affect aquatic species could be beneficial in understanding and mitigating these issues.
References


precocious maturation of early-rearing Atlantic salmon *Salmo sala*. 
*Aquaculture Research, 49*(2), 801–808. https://doi.org/10.1111/are.13511

*Hydrological Sciences Journal, 54*(1), 101–123. 
https://doi.org/10.1623/hysj.54.1.101

THE ROYCEAN COMMUNAL IDEAL IN T. S. ELIOT'S "THE LOVE SONG OF J. ALFRED PRUFROCK"

MICAH GILL, HARDING UNIVERSITY
MENTOR: NICK BOONE

Abstract

T. S. Eliot’s poem “The Love Song of J. Alfred Prufrock” has often been read as a thoroughly cynical assessment of the modern individual and community. This totalizing portrayal misses the poem’s ability to serve as a foil to a more ideal community. If Eliot is so ill-disposed to the timid, fragile Prufrock and his shallow, ostentatious social network, then Eliot must believe in the essence of a better community. We may look to one of Eliot’s intellectual mentors, Josiah Royce, for his philosophy of community, which stands as an ideal inverse in relation to the Prufrockian community. After examining Royce’s thought and reviewing the scholarly dialogue on this topic, this article contrasts the inferior Prufrockian community with the Roycean community, its ideal counterpart, on their respective communication styles, quality of interpersonal relationships, and conceptions of selfhood. These comparisons elucidate Royce’s vision for social living and demonstrate how “The Love Song of J. Alfred Prufrock” affirms the possibility for flourishing human community.

In T. S. Eliot’s poem “The Love Song of J. Alfred Prufrock,” the titular protagonist imagines a series of social interactions that leave him feeling empty and misunderstood. Prufrock describes these disappointing interchanges within the structure of a winding dramatic monologue. This technique gives the reader insight into his fearful, reactive psychological state, which is particularly evident when he is confronted with the possibility of interpersonal relationship. The hypothetical conversations that Prufrock describes always fail to create meaningful mutual understanding and connection, suggesting that the protagonist’s actual relationships, which would shape his imagination, are similarly unfulfilling.

Eliot foregrounds the difficulty of social interaction in the poem not only to provide a dialogic structure in which to voice his philosophical ideas but also to comment on the topic of community (i.e., a network of individuals who share lived experience through interpersonal relationship). Several factors directed Eliot to consider the potential for a functional community at the time of writing the poem in 1911, including the sociopolitical climate of the West leading up to World War I and the contemporary academic discourse influenced by philosophers such as Josiah
Royce. Eliot’s understanding of community developed and changed dramatically throughout his career, so any of Eliot’s individual depictions of community should not be taken as representative of his whole thought. Arriving at the beginning of his poetic career as his first published poem, though, “The Love Song of J. Alfred Prufrock” constitutes an important landmark in understanding the evolution of Eliot’s perspective on community.

In his internal musings, Prufrock constitutes what we may call “the Prufrockian community,” an imaginary social and relational structure that includes himself and other individuals he refers to in the poem. This society is characterized by fractured communication, relationship, and selfhood. Prufrock’s jaded vision of what a community can be, however, does not necessarily reflect Eliot’s view. The poet does not condemn all community to the same broken fate but instead casts the broken Prufrockian community as a negative manifestation from which the reader may infer a positive, ideal inverse. This inverse is a flourishing community that resembles and is clarified by Josiah Royce’s vision of an interpretive community, one where there is honest communication, meaningful connection, and communally integrated selfhood.

Eliot’s intellectual kinship to Josiah Royce figures prominently into his views on community and its portrayal in the “The Love Song of J. Alfred Prufrock.” At the time Eliot was writing the poem, Royce was teaching at Harvard and developing a pragmatic, idealistic philosophy that he would use to formulate a framework for community. After completing undergraduate studies at Harvard in 1910, Eliot applied himself to a year of intellectual growth in Paris before returning to Harvard in 1911 to begin doctoral studies. After deliberating upon a dissertation topic, Eliot eventually decided to write about and research the epistemology of F. H. Bradley, and Royce was appointed as Eliot’s dissertation supervisor. Their close philosophical ties were deepened when Eliot involved himself with Royce’s 1913 seminar on interpretation entitled “A Comparative Study of Various Types of Scientific Method,” and this experience shaped Eliot profoundly (Gray et. al). Though Eliot had written “The Love Song of J. Alfred Prufrock” shortly before commencing his studies under Royce, he would have already been very familiar with Royce’s work from the time they had previously shared at Harvard.

The community that Prufrock envisions in the poem is made more intelligible through contrast with Royce’s ideal interpretive community. M. L. Briody defines the Roycean conception of community as “a reality constituted by unique individual selves who share a common history and/or aim” (226), and Royce’s ideal version of this community, which he calls “the Beloved Community,” is unified by “loyalty, which is the love of a self for an [sic] united community” and shares the aim of helping its members to understand the world and thrive in it (Royce, War and Insurance 34). Another key aspect of the ideal Roycean community is that it is interpretive. Royce’s epistemology is summarized as follows:
Knowledge is not at bottom merely the accurate and complete perception of an object, as empiricism would have it. Nor is it the accurate and complete conception of an idea, as rationalism maintains. Knowledge is instead a process of interpretation: the true idea selects, emphasizes, and re-presents those aspects of the object that will be meaningfully fulfilled in subsequent experience. (Parker and Pratt)

In an interpretive community, the members approximate an increasingly accurate understanding of their being and world through collaboratively interpreting meaning, experiences, and objects. Parker and Pratt quote Royce in suggesting that, in an ideal Roycean community, this objective for understanding through interpretation is completely fulfilled: “If the interpretation is a reality, and if it truly interprets the whole of reality, then the community reaches its goal.” Through contrasting the Prufrockian community to Royce’s Beloved Community, readers will develop a fuller picture of the latter. This insight is unlocked through acknowledging the importance of community as a theme in the poem and maintaining an awareness of the degree of Royce’s influence on Eliot.

The scholarly discussion on the ideal community in “The Love Song of J. Alfred Prufrock” can be advanced by demonstrating that the poem advocates for a Roycean approach to community and elucidates the nature of that approach. James C. Haba originates the discussion by showing how Prufrock’s distorted perception of himself and his failure to observe the humanity in other people precludes his ability to enter meaningful, communal connection (Haba). Haba identifies many of Prufrock’s misunderstandings about community, and Robert McNamara later extends these insights by showing how Prufrock’s distorted view of self drives much of his alienation (McNamara). McNamara argues that Prufrock’s narcissism, which reflects the fragmentation of the modern self, drives Prufrock away from real people and relegates him to merely abstract connection. Abstraction is the only place where Prufrock can cohere a self that he desires while maintaining a rigid individuality.

Royce’s philosophy would condemn the Prufrockian self as explained by McNamara, which is part of the reason why Charles Anthony Earls introduces the Roycean framework of community into the dialogue surrounding the poem. Earls applies Royce’s concept of the interpretive community as an antidote to the communicative and relational issues of the Prufrockian community (Earls). Earls then redirects his application of Royce toward refuting Frank Lentricchia’s contentions against Roycean communal paradigms (Lentricchia). Kate McLoughlin then extends Earls’s connection between the poem and the Roycean communal paradigm through a more focused literary and critical analysis. In this analysis, she leverages Royce’s framework to identify some of the distorted aspects of and challenges within the Prufrockian community (McLoughlin). More thoroughly than in Earls’s approach, McLoughlin demonstrates how the Prufrockian community’s
failure as a Roycean interpretive community paralyzes Prufrock in fear, isolation, confusion, and inaction, then suggests that this failure reveals an underlying cynicism in the poem about the potential for any true community to arise.

There is now the opportunity to synthesize and advance the discussion. Earls’s work sees a validation of the Roycean communal framework in “The Love Song of J. Alfred Prufrock,” but his treatment of the poem was brief and primarily served as a tool to combat Lentricchia. His strong ideas pairing with a merely tangential focus resulted in an approach that was potent but lacked the breadth and depth of a thorough critical analysis. McLoughlin’s more extensive contribution leverages Earls’s lens but fails to acknowledge the potential for a positive affirmation of an ideal community to arise out of the poem. McLoughlin reads the poem as an admission of the impossibility of true interpretive community, when instead, Eliot implies that the reader should seek beyond the failed communal attempts of the poem to envision an inverse approximating Royce’s Beloved Community. I will attempt a synthesis of these viewpoints by demonstrating how, through portraying the Prufrockian community as an inverse, the poem clarifies Eliot’s conception of the Roycean community. In addition, I will attempt to incorporate previously undiscussed examples from the text that enhance the juxtaposition between the Prufrockian community and Royce’s Beloved Community; this is done so we may more fully realize the poem’s power to elucidate the nature and characteristics of the ideal community.

To contrast the nature of the Roycean community with its Prufrockian foil, we will first examine an individual quality of the Roycean community before contemplating the inverse quality manifested in the poem, starting with the theme of communication. It is evident, even etymologically, that communication fosters community; the Roycean community, though, considers communication as not merely socially preferable but also epistemologically vital. In Royce’s interpretive epistemology, a Community of Interpretation “is the totality of all those minds capable of representing aspects of Being to one another or to their future selves” (Parker and Pratt). Representing Being to each other equips the members to better understand the world around them and to operate effectively in it. Because the interpretive process helps others in the community to live well, Royce proclaims that “no one who loves mankind can find a worthier and more significant way to express his love than by increasing and expressing among men the Will to Interpret” (The Problem of Christianity 218). In addition, the act of interpretation expresses love not only for the individual members of the community but also for community as such—“the purest forms of love for community possible” (Royce, The Problem of Christianity 218)—by acknowledging community’s telos as an epistemological tool.

Communication in the Prufrockian community, though, is concealing and fractured. Instead of a community of people honestly and openly interpreting reality with each other, Prufrock can imagine only a community that is accustomed to trivial,
surface-level dialogue shrouded with pretense: “Among the porcelain, among some talk of you and me” (line 89), interpretive communication is unable to flourish. The community members’ discussions are often little more than “Talking of Michelangelo” (line 14), bantering about foreign cultures as a platform for ostentation instead of shared appreciation. Eric Sigg notes that “Eliot’s earliest poems quietly register the American tendency to associate culture with what is foreign,” and that Eliot extends this American impulse to the Prufrockian community specifically as a pattern of pretense when discussing foreign culture (19). McLoughlin affirms that the Prufrockian “conversations about art and novels do not constitute ‘the historical sense’ but are representative of the sort of ‘knowledge’ that Eliot noted could be ‘put into a useful shape for examinations, drawing-rooms, or the still more pretentious modes of publicity’” (55). Such is the duplicitous, manipulative nature of communication that characterizes the Prufrockian community.

The Prufrockian community’s conversation defaults to triviality, demonstrating how Prufrock has no understanding of a community in which members help each other interpret that which is important in life. He questions whether he would ever be allowed to “have squeezed the universe into a ball / To roll it towards some overwhelming question” (lines 92–93) in any of his interpersonal relationships. The “overwhelming question[s]” are exactly what a healthy interpretive community should address, yet the “tea and cakes and ices” (line 79) constitute an impenetrable barrier of banal pleasantries that Prufrock thinks would frustrate any of his attempts to “force the moment to its crisis” (line 80). Prufrock imagines the woman in the poem giving up in the face of the community’s nearly insurmountable obstacles to pure, honest, interpretive communication with a lament, “That is not what I meant at all; / That is not it, at all” (lines 97–98). Prufrock encapsulates his frustration with such fragmented communication when exclaiming, “It is impossible to say just what I mean!” (line 104). McLoughlin summates the communicative quandary: “Here are no communities of memory or hope, no consensual epistemology, but irredeemable inattention, misunderstanding and crossed purposes” (55). McLoughlin’s bleak diagnosis captures the inability of the Prufrockian community to corporately articulate and interpret what is true, significant, and beneficial.

By portraying vapid communication within the Prufrockian community, Eliot enables the reader to more clearly envision a hypothetical, Roycean inverse. This ideal community, clarified through contrast with the Prufrockian, would communicate transparently and honestly about all things. Its members’ conversational scope would not only encompass simple matters but also broach the difficult and profound questions of life in courageous interpretive acts. Earls sees how this interpretive community would remedy the Prufrockian communicative divide, as “members of this community would be united by their ambition to have Prufrock fully comprehend what his companion means, and in turn to have her understand Prufrock’s desire” (131). This pure communication—like “the mermaids singing, each to each” (line
—is yet unattained by the Prufrockian community: “I do not think that they will sing to me” (line 125), Prufrock sighs in recognition of the interpretive deficit within his envisioned community.

The Roycean community concerns itself not only with deep, honest communication but also with similarly meaningful forms of relationship. The significant communication that occurs in an interpretive community is soul-baring and requires a type of vulnerability that strengthens interpersonal connection, as made evident by “current intimacy process models [that] suggest that vulnerable disclosures promote partner responsiveness and therefore increase intimacy” (Khalifian and Barry). Vulnerability increases intimacy through such means as “sharing personal information,” which “is a way of communicating trust and desire to share one’s self with another” (Khalifian and Barry). The interpersonal relationships in a Roycean community that encourages such communication will therefore be close and deep, a reality that Royce envisions when describing a community in which “we know even as we are known” (The Hope of the Great Community 35). In this light, Earls sees how, if Prufrock practiced less isolation and more vulnerability, he would be better able to understand the possibility of a healthy community with robust interpretive capacity: “If the attempt at mutual self-disclosure is successful, then the strategy leading to its success is likely repeated” (131). Individual members of the Roycean community, though, are not required to homogenize if they want to attain this depth of relationship. Instead of calling members of the Beloved Community to sacrifice their uniqueness, Royce declares that “a community does not become one, in the sense of my definition, by virtue of any reduction or melting of these various selves into a single merely present self, or into a mass of passing experience” (The Problem of Christianity 67). In a community replete with vulnerable, interpretive communication and in which uniqueness is celebrated, members will naturally connect through articulating the diverse, significant aspects of their shared lived experience.

The Roycean community’s relational intimacy contrasts with the shallow connection between members of the Prufrockian community. In Prufrock’s view of the world, the most prominent romantic venues are not chapels that feature the exchange of covenantal vows but are instead “one-night cheap hotels” (line 6). Noncommittal, licentious flings are the predominant mode of Prufrockian connection. Drawing rooms and parties also house transitory relationships, as “in the room the women come and go” (line 13). Prufrock describes these individuals as wandering in and out of community interactions without establishing any ties deeper than proximity. Even when members of this community are together, real connection is scarcely possible when Prufrock and his imagined interlocutors “prepare a face to meet the faces that you meet” (line 27), creating and acting out a facade instead of opening themselves to meaningful relationship.
Prufrock also projects his tendency to create inaccurate caricatures of others, which further inhibits authentic relationships from forming among members of his imagined community. Instead of acknowledging the individual complexity of their fellow humans or seeing the possibility for building a stronger relationship with them, Prufrockian community members assume that they have already plumbed the depths of relational potential and have come up empty. Prufrock does not see others’ complex motives and personhood, so the community he envisions likewise sees each member as little more than a “patient etherized upon a table” (line 3), fit for surgically dissecting and critiquing but not for loving or knowing. Later in the poem, Prufrock rejects the possibility of authentic conviviality; his community members view each other less as friends and more as scientists who mechanically examine insects would: “And I have known the eyes already, known them all– / The eyes that fix you in a formulated phrase” (lines 55–56). Prufrock can see that his approach to relationship leads to shallow dead-ends; the issue is that, in his own narcissism and anxiety, he cannot imagine a form of companionship or community that would reciprocate a step toward the type of vibrant relationships we see in Royce’s community. The reader can hear Prufrock’s exasperated tone in facing this interpersonal quandary, where he remains hesitant to reveal any personal information that could spark meaningful relationship:

When I am pinned and wriggling on the wall,
Then how should I begin
To spit out all the butt-ends of my days and ways?
And how should I presume? (lines 58–61)

Prufrock cannot imagine others caring for authentic relationship, so he chooses a reticence that relegates him to surface-level connection rather than having others scoff at his vulnerability.

Prufrock is filled with an existential dread when recognizing the paucity of meaningful interpersonal relationships within his life, a void that he also projects onto the Prufrockian community. He reflects, “I have measured out my life with coffee spoons” (line 51), signifying that his social experience has been characterized by empty interactions instead of the meaningful events that are more common within a flourishing community. Interpersonal milestones such as marriages, funerals, childbirths, and adventures with friends, mark the lives of relationally invested individuals operating in an ideal community. Even commonplace activities, such as sharing meals with loved ones, become much more significant and contributive toward a flourishing human experience when they are conditioned by deep relationship.

When Prufrock envisions community, though, he sees only idle, isolated activity, “the smoke that rises from the pipes / Of lonely men in shirt-sleeves, leaning out of windows” (lines 71–72). The image suggests a vapid, somber existence, in which
solitary musings stand in for true connection. Perhaps a first step toward relationship would occur if they came down from their windows and smoked together on the street, vulnerably admitting their desire for companionship and having that desire reciprocated. Instead, as Prufrock meanders alone “at dusk through narrow streets” (line 70), he feels as anonymous and unseen as a crab on the desolate ocean floor: “I should have been a pair of ragged claws / Scuttling across the floors of silent seas” (lines 73–74). The sense of solitude perpetuates Prufrock’s anxiety when he considers death: “And I have seen the eternal Footman hold my coat, and snicker” (line 85). He feels as though death itself mocks his isolated state and condemns any dream of finding more substantial relationship before the grave.

Reflecting on the Prufrockian community’s relational void enables us to perceive the obstacles preventing flourishing relationships and the significant experiences they provide. By vulnerably sharing their authentic selves, investing in consistent rather than transitory relationships, and refusing to create shallow caricatures of their peers, community members will be closer to attaining that Beloved Community that enriches their lives.

In addition to advancing meaningful communicative and relational practices, the Roycean community also fosters a specific sense of self. As mentioned, Royce did not envision a complete envelopment of the individual self when adjoining to the Beloved Community. In his view, the self is not compromised but fulfilled in the ideal community. Briody elaborates that in Royce’s view, “Through each man’s effort toward forming community through interpretation, his own individuality is enhanced rather than diminished. He becomes more uniquely himself precisely through his communal relationships” (229). The self and the community integrate and exist symbiotically instead of dissolving into each other. The individual maintains their uniqueness while contributing to and benefiting from the interpretive, relational, and edifying aspects of membership in the Beloved Community. If “it is from his own unrepeatable individuality that a man contributes to the greater wealth that is community” (Briody 225), then uniqueness is not seen as a barrier to communal unity but as its supplement.

Individuals’ identities integrate into Roycean community primarily through sharing “a common past and/or a common future. In regard to their past, the group constitutes a community of memory; in regard to the future, a community of hope” (Briody 226). The community members must share core values for understanding the past if they narrativize their history similarly. If they have a common hope for the future, then they are allied under beliefs that inform what to strive for and how to attain their mutual goals. Such a fundamental point of connection empowers members to “act as a community as well as individuals. In so acting, they achieve a personal reality over and above their isolated individualities” (Briody 226). It is when living out
of unified values and beliefs in a communal context that the individual self and the community are most fully integrated.

The Prufrockian community distorts Royce’s vision of an authentic, integrated self and community by enforcing a surface-level homogeneity, under which there are no shared values. Prufrock highlights the monotony of his envisioned community's polite drawing-room interchanges to describe the disingenuous roots of their behavioral unity. He sighs, in an exasperated and fatigued tone, “For I have known them all already, known them all—/ Have known the evenings, mornings, afternoons” (lines 49–50), “I know the voices dying with a dying fall / Beneath the music from a farther room” (lines 52–53); “I have known the eyes already, known them all” (line 55), and “I have known the arms already, known them all” (line 62). Prufrock does not have faith in the possibility of a community that properly celebrates that which is different and “other.” Instead of embodying the Roycean ideal in that way, the Prufrockian community militates a behavioral uniformity that belies disparate motives and clashing values beneath a manicured surface.

We may perceive this deep dissonance by examining Prufrock’s inner turmoil when he considers seemingly harmless social contexts. Maintaining homogeneity distresses Prufrock, as doing so would force him to compromise his identity and conform instead of expressing who he truly is, yet he has no point of reference for a community that would welcome him without changing him. His monologic musings on relationship, forming the basis of the Prufrockian community, express this anguish in his repeated refrain “Do I dare?” and in his indecision on whether to enter stressful social interactions by allowing “Time to turn back and descend the stair” (lines 38–39).

Moreover, Prufrock does not view social interactions as fluid, natural, and enjoyable, as they often are in a community where people love each other and share values. Prufrock instead anticipates “time yet for a hundred indecisions, / And for a hundred visions and revisions” throughout when conversing with others (lines 32–33). Prufrock’s psychological and emotional volatility manifests in strained, manufactured rapport; he can imagine little else among people who lack a true shared identity to connect through. Amidst the social scrutiny arising from fragmented identities, he even views such trivial breaches of conduct as parting his hair a different way and eating a peach as daring endeavors (line 122), and he attaches cosmological significance to minor impropriety: “Do I dare / Disturb the universe?” (lines 45–46). Prufrock feels compelled to perform an insincere self if he wants to qualify for community. He risks social rejection if he fails this performance on the stage of the drawing room, and this possibility of alienation amplifies his anxiety.

While the Prufrockian community represents a distorted view of communication, relationship, and self in community, the end of the poem suggests Prufrock’s ability to change in closer resemblance to the Roycean ideal. McNamara
notes how Prufrock is not rendered “as a unified and coherent self, but rather as a figure paralyzed by his narcissistic investments” (372), but near the poem’s conclusion, Prufrock begins to shed his narcissism and develop a view of self that is more in keeping with membership in a flourishing community:

No! I am not Prince Hamlet, nor was meant to be;  
Am an attendant lord, one that will do  
To swell a progress, start a scene or two,  
Advise the prince; no doubt, an easy tool,  
Deferential, glad to be of use. (lines 111–115)

Though this monologue has often been read as Prufrock languishing in his insecurities, we may just as reasonably read it as Prufrock integrating his view of himself into community. He is casting off his self-centeredness that previously convinced him of the need to be the beloved star of the drawing room. In a Roycean community with shared hope and future, members recognize their individual roles in contributing to that future and acknowledge that not everyone is given a lead role in driving the plot forward. Prufrock is starting to realize the significance of others around him, causing him to reflect on whether he needs to be the hero for his community to flourish.

It is this realization, where Prufrock begins to break through his narcissism, that Haba sees as completed in the final line of the poem. Haba explicates “Till human voices wake us, and we drown” (line 131) as Prufrock realizing the value and humanity of others in his actual, not imagined, community. Prufrock’s realization mirrors Christian baptism in that he “drowns” his old self and previous isolation and resurrects with a new self that is integrated into an ideal community resembling the Church (Haba 56–59). Prufrock wakes up from that monologic dream in which he paints a failed vision of community where only his voice resounds. He now distinguishes the actual other voices around him, no longer muddling them with the imagined Prufrockian community, thus opening the possibility for true relationship. Despite the difficulty of realizing Royce’s ideal community, Eliot portrays the potential for positive communal change at the end of the poem, challenging readers and Prufrock himself to attain the Beloved Community.

Though the beginning period of Eliot’s career, in which he wrote “The Love Song of J. Alfred Prufrock,” is perhaps his most cynical, we still find prominent features of hope for the self, the community, and the integration of the two in the poem. Considering that Royce compared his theorized Beloved Community to the Christian Kingdom of Heaven, it is unsurprising that, after a “long meditated and perhaps even long deferred” process (Kearns 88), Eliot converted to Christianity. Perhaps his vision for an ideal community, portrayed inversely in “The Love Song of J. Alfred Prufrock,” led him to pursue that fellowship in the religious venue where he deemed it most likely to be found. Regardless, the Prufrockian community’s distorted practices of
communication, relationship, and selfhood serve as a foil to better perceive the characteristics, worthiness, and partial attainability of the ideal, Roycean community.
Works Cited


YAYOI KUSAMA’S DEPICTIONS OF MENTAL HEALTH

ANNA ISABEL DAVIS, TRANSYLVANIA UNIVERSITY
MENTOR: EMILY GOODMAN

Abstract

This paper examines Yayoi Kusama’s usage of art as a coping mechanism for her mental illness, particularly through her pumpkins and “nets” motifs. It involves a close analysis of three artworks to emphasize this motif: her first depiction of pumpkins; “nets”; and, finally, her combination of both pumpkins and “nets” into one artwork. Her theme of repetitions as meditative force is evident in each of these works. As such, this paper posits that the goal of her art is not to heal from her mental illness but to better cope and adapt. This research draws on various primary sources, including Yayoi Kusama’s autobiography, Infinity Net, and intertwines three artworks that have not yet been discussed together. It also has broader implications for other artists because Kusama paved the way for a more open discussion about mental health in contrast to the trope of the suffering artist.

For an aspiring artist like myself, to triumph over an unjust environment is to triumph over the pain of feeling cornered and trapped. I see it as a trial or test attendant upon having been born a human being, which is why I continue to fight with every fibre of my being. This is my own peculiar karma and destiny in this world.

Yayoi Kusama

In more recent decades, the impact of mental health on an individual’s life and career has moved more to the forefront. Given the relevance of this discussion, it is important to see an artist in terms of not just their mental health struggles but also their coping mechanisms. Generally, when we think of the mental health of artists, we first think of the narrative of the tortured artist. Artists such as Vincent van Gogh and Mark Rothko are often lauded and their struggle considered “beautiful” and “necessary” for the creation of their works. Although discussing the reality of artists’ struggles is important, the narrative of artists who have continued to fight their battles with mental illness is often overlooked. Artists such as Tracey Emin (Figure A1) and Yayoi Kusama (Figure A2) have found healthier ways to cope and survive with their mental illnesses, all while including in their art practices the truth of their struggles and their methods of management.
In this paper, I will be examining how Yayoi Kusama, an artist who has been and continues to be very open about her struggles with her mental health, has used her practice to cope with her obsessions, compulsions, and hallucinations. In particular, I will analyze specific themes in Yayoi Kusama’s art, especially her depictions of pumpkins as representations of her hallucinations, her use of nets as a form of meditation, and the instances when she combines the two as markers of her mental health journey.

Yayoi Kusama was born March 22, 1929, in Matsumoto, Nagano, Japan, to a family of merchants (Figure A3). She was the youngest of four children and demonstrated an interest in art from an early age. Her mother discouraged Kusama from painting and pushed for her to be a traditional housewife, but her father was very supportive of her art and continuously bought her supplies. While the trauma of World War II and various family issues made her artistic training more difficult, at the age of twenty-nine, Kusama was able to formally pursue art and train at the Kyoto City Specialist School of Arts.

Kyoto was not far enough away from her family, and Kusama continued to feel stifled by her mother’s expectations that she adhere to the demands of a traditional and conservative Japanese woman. In her autobiography, Yayoi Kusama talks about leaving Japan:

Staying in Japan was out of the question. My parents, the house, the land, the shackles, the conventions, the prejudice. . . . For art like mine—art that does battle the boundary between life and death, questioning what we are and what it means to live and die—this country was too small, too servile, too feudalistic, and too scornful of women. My art needed a more unlimited freedom, and a wider world.

After eight years of attempting to convince her family to allow her to leave Japan and study art abroad, Yayoi Kusama finally made her artistic journey to America on November 17, 1957. In 1958, she moved to New York and began to freely practice her art, continuing to do so until 1973, at which point she returned to Japan to seek treatment for ongoing mental health struggles. She has remained in Japan ever since.

Although Kusama started seeking treatment in 1973, her mental health struggles began long before then. Kusama had a rather traumatic childhood. Her parents had a strained relationship, and Kusama’s mother was physically and verbally abusive toward her, including forcing her to spy on her father when her mother suspected him of womanizing. Her mother also restricted Kusama’s access to art supplies and initially prevented her from traveling abroad to pursue her career as an artist. This led to Kusama feeling like she had to escape her home and family, although
running away from these experiences proved to be inadequate for remedying the struggles she felt as a result of them.

In addition to Kusama's multiple adverse childhood experiences in her home life, World War II ruined any remaining semblance of her childhood. The war left a mark on Japan, which lost “approximately three percent of the Japanese population, [or] more than 2 million people” by the war's end in 1945. For civilians, life in Japan became very difficult: “During the war, more than 1 million Japanese children were evacuated from cities to the relative safety of the countryside, splitting up families in the process. Rationing and food shortages were rampant throughout the country, beginning as early as 1940.” In addition to evacuations, millions of women and children, including Kusama, were enlisted to work in factories to help the war effort. At this time, Kusama—who was about 11 years old—began working twelve-hour days in a factory, making uniforms and parachutes for the war with other kids her age. According to Judith Rodenbeck, “The people of Japan were living under starvation conditions, sharing recipes for insects and grass while being bombarded with martial propaganda from their own government and with a rain of fire from Allied air forces.” Despite these excruciatingly long days, Kusama continued to practice art on the side, proving how important it was to her.

In addition to these traumas, in 1936, at the age of seven, Kusama began to experience hallucinations. It is important to note that her hallucinations may have occurred without the impacts of the war or family trauma, but the existence of these events likely exacerbated her symptoms. According to Aleandra Munroe, Kusama was diagnosed with “an obsessive-compulsive and hysterical condition . . . attributes her manic obsession with themes of repetition, aggregation, and accumulation to a recurring hallucination.” Her hallucinations included an instance when flowers from the tablecloth appeared to cover the entire room in which she was standing. More frequently, her hallucinations have consisted of flowers and flashes of light, but, of course, the aspect of her hallucinations that affected her the most was the pumpkins that spoke to her, as proven by the multitudes of pumpkins in her art.

These symptoms continue to plague Kusama to this very day. In her autobiography, Infinity Net, Kusama says, “I fight pain, anxiety, and fear every day, and the only method I have found that relieves my illness is to keep creating art. . . . I followed the thread of art and somehow discovered a path that would allow me to live. If I had not found that path, I am sure I would have committed suicide early on, unable to bear the situation in which I found myself.” Kusama has made it clear throughout her life that art has been a constant therapeutic outlet without which she would not be able to manage her mental health.

It is important to note that Kusama’s goal is never to be completely healed from these issues, as that is impossible. Her only goal is to adapt, live, and work around her mental health. Her artistic practice has continued to develop, as have her coping
mechanisms, and the various themes in her work demonstrate the progress she has made in both arenas. In her autobiography, Kusama discusses managing her mental health using art: “The only way for me to elude these furtive apparitions is to recreate them visually with paint, pen, or pencil in attempt to decipher what they are; to gain control over them by remembering drawing each one that flashes through the haze.” Kusama believed that, by depicting what she saw, she was able to better manage and analyze her hallucinations.

In 1948, Kusama created her first representations derived from one of her hallucinations, that we know of (Figure A4). This painting depicts three almost white pumpkins on a checkered black and brown background. The background depicts the repetition that would become more prominent in her work throughout her career. The pumpkin on the far left is mostly white with some orange, is cracked on the top, and has a stem. The middle pumpkin is more orange than the other two but still has some white, with a stem that sticks straight up. The final pumpkin is depicted with cracks on the top and has some orange, though it is mostly white, with no stem, making it look more like an onion than a pumpkin. These pumpkins are far different than the animated ones we are used to seeing in her more contemporary works, but they still depict the cannon of repetition, given that she includes multiple pumpkins in this same painting.

Throughout Kusama’s life, pumpkins would remain a constant, no matter her mental state. She writes in her autobiography, “During my time in Kyoto I diligently painted pumpkins, which in later years would become an important theme in my art.” She notes, “It seems that pumpkins do not inspire much respect. But I was enchanted by their charming and winsome form. What appealed to me most was the pumpkin’s generous unpretentiousness. That and its solid spiritual balance.” She also recalls having consumed the vegetable endlessly to the point of nausea in her childhood years during and after the war; in spite of this, she retains a fond attachment to its organic bulbous form.

Additionally, the pumpkins are relevant because, she has said, they have repeatedly appeared in her hallucinations. In her autobiography, Kusama talks about the first time she saw a pumpkin and the hallucination that went along with it: “I parted a row of zinnias and reached in to pluck the pumpkin from its vine. It immediately began speaking to me in a most animated manner.” Her painting therefore acts as a form of therapy to process her traumas; thus, when she depicts them in her art, she is both making something that reminds her of the tumult of her upbringing and making something that appeared in her mind appear in the real world and to others. Bringing to life in a painting something that only she could see brought Kusama comfort, not only to try to control the vision, but also to share it with others.

In addition to depicting pumpkins, her first painting includes the geometric repetition that is so prominent in much of her work, through the checkered
background. As her practice developed, particularly after she moved to New York in 1958, Kusama began working on her “net motif.” According to Midori Yamamura’s review of *Infinity Net*, “She repeatedly has said that she suffers from hallucinatory visions where she sees ‘nets’ and ‘dots’ that profoundly affect her artist development and manifest themselves in her visual vocabulary.” Kusama’s “nets” (Figure A5) are simply a single brushstroke repeated over and over again, done in contrast to another color. She has created many different versions (Figure A6), but all of them hold the same basics of two colors, while some have more defined shapes and others more intense variation of color.

For Kusama, the “nets” are a deliberate act of meditation. She began working on them as her mental health took a rapid decline. She filled canvas after canvas with “nets.” According to her autobiography,

> I often suffered episodes of severe neurosis. I would cover a canvas with nets, then continue painting them on the table, on the floor, and finally on my own body. As I repeated this process over and over again, the nets began to expand to infinity. I forgot about myself as they enveloped me, clinging to my arms and legs and clothes and filling the entire room.

This process continued for as long as she needed. The “nets” acted as a method to calm her and make her stop thinking of the panic she was feeling. This use of repetition as a form of self-soothing in her work is helpful to understand how her mental health has informed her process, especially since she includes it within almost all of her works.

Although her “nets” have helped to calm her in the moment, they did not prove to be a long-term solution for her to cope with her mental struggles. Throughout her time in New York, her mental health began to deteriorate. According to Midori Yamamura, “Donald Judd (Kusama’s upstairs neighbor) recalled that she had suddenly become ‘very paranoid about the New York art situation.’ Another neighbor . . . remembers that she suddenly became obsessed with the thought that her ideas might be appropriated, which compelled her to close all the curtains on the windows of her loft.” During this time, “Kusama almost surely suffered from anxiety, neurosis, and to calm herself . . . she mentions that she took Doriden (a prescription ‘minor tranquilizer’).” As these struggles increased, she made attempts to manage her symptoms both through her art and through some self-medication before ultimately realizing that she needed professional help.

In 1977, a few years after the death of her father, Kusama’s mental health deteriorated to the point where she was no longer able to care for herself. Kusama checked herself into a hospital for the mentally ill in Tokyo, where she continues to voluntarily reside. For her to check herself into a mental hospital in the 1970s means her mental health was deeply unmanageable, because seeking treatment for mental health was at the time, and somewhat remains, stigmatized. In the early 1960s, Japan
had begun moving in a more progressive direction with regard to the treatment of mental illness, having repealed acts that kept mentally ill people out of society and moving toward mental health care becoming more community-based. In 1964, however, the United States ambassador to Japan, Edwin O. Reischauer, was stabbed by a nineteen-year-old who was having a breakdown, causing a regression in policies toward mental health treatment and the revision of the Mental Hygiene Law in 1965. According to Tomoko Kanata, Professor of Intercultural Studies at Kobe College Japan, “It was quite obvious that its main concern was to safeguard the public from people with ‘dangerous mental disorders.’ In order to do so, people considered to be a danger to the public because of their mental condition needed to be cared for and monitored in communities as well as in hospitals so that they would not be able to harm others.” Given that people with mental illness were considered dangerous in Japan at that moment, Kusama’s decision to return home and receive in-patient care was a testament to the profoundness of her mental health struggles and her determination to get help.

None of this stopped her from continuing her artistic journey. At the inpatient facility, she has continued to create art and even wrote and published her autobiography. She is able to maintain and work in her own studio, and she has created new works that have gone to galleries, exhibits, and museums across the world. She has created her exceptionally famous infinity rooms, and a museum dedicated to her opened in 2014. This has worked so well that she still lives in the facility today, at 94, making it so far a total of forty-six years she has had the same lifestyle.

In the many years that Kusama has been receiving treatment, she has continued to make works that involve elements of her struggles with her mental health as well as techniques that she has used to cope with them. Pumpkins remain a prominent theme in her work, but in the majority of Kusama’s depictions of them, she overlaps her pumpkins with her “nets” (Figure A7). This is true for both her paintings and sculptures. Her more recent depictions of pumpkins combined with “nets” are very stylized and in some instances almost abstract, which is starkly different from her earlier portrayals of pumpkins. In these works, the “nets” appear more as a series of polka dots, as opposed to a web pattern. Additionally, her earlier depictions of pumpkins were muted in color, whereas the pumpkins combined with the “nets” use brighter colors. These transformations in her portrayals of pumpkins and “nets” reflect the progress of her artistic development and in her struggles with mental health.

Together, they not only combine common motifs in her work but also, specifically, align the object of her hallucinations with the meditative practice she has developed to cope with them. In creating this, she transforms her pumpkins from a symbol of her disorder into something of peace and calm. As such, “The pumpkin stands as a symbol of triumph for the artist’s personal as well as artistic rebirth,
representing a mediation of the artist’s psychiatric illness that went hand-in-hand with the ever-increasing sophistication, dexterity and creativity of her creations.” That she has become known for these bright, spotted pumpkins demonstrates that she has transformed her struggle into meditation and that meditation into a career.

Yayoi Kusama’s portrayals of her hallucinations and her methods of coping offer a counternarrative to the myth of the tortured artist, particularly that of figures like Vincent van Gogh. Van Gogh and his work are generally looked at through the lens of a beautiful tragedy because of the way his life ended, rather than through a lens recognizing the truth of his keen eye and appreciation of beauty. The myth of the tortured artist is deeply toxic because it glorifies suffering as central to the creation of great art and in turn promotes the idea that seeking treatment or getting help will diminish that work. All of this works to elevate suffering over survival.

The myth of the tortured artist is so heavily intertwined with van Gogh that even Kusama has commented on it. At the end of her autobiography, she writes,

Many people seem to imagine that Vincent van Gogh must have been great because his paintings now fetch enormously high prices or because he was mentally ill. But such people have not really seen van Gogh. . . . My view is that in spite of whatever illness he may have had, van Gogh’s art overflows with humanity, tenacious beauty, and the search for truth. His real greatness lies in these qualities, and in his fiery and passionate approach to life.

Kusama thus acknowledges the mythologizing of van Gogh’s struggles but notes that his work contains many other points of connection, ones that are often overlooked in favor of focusing on how his work reflects his suffering.

In opening up about her mental health, Kusama has likely shifted the discussion of mental health within the art world. Through the repetition and meditation of Kusama’s “nets” and pumpkins, she redefines her mental health from a setback into a strength. In creating pumpkins and “nets,” she brings her hallucinations to the real world while layering on top the meditative process that keeps them at bay. In doing this, she has changed the narrative in which we see and discuss mental health struggles, particularly when represented in an artist’s work.
References


Appendix

Figure A1. Tracey Emin, It was all too Much (2018)

Figure A2. Yayoi Kusama, the Artist in Her Custom Dots Obsession BMW Mini Car (2004)
Figure A3. Kusama with Her Family (approx. 1930s)

Figure A4. First Known Painting by Yayoi Kusama (1948)

Figure A5. Yayoi Kusama, *NO. F* (1959)
Figure A6. Yayoi Kusama, *NO. Red B* (1960)

Figure A7. Installation View of Yayoi Kusama: *THE MOVING MOMENT WHEN I WENT TO THE UNIVERSE*, a Victoria Miro Gallery (2018)
PARTAKE OR BE DEPARTED: INSULARITY OF SETTING WITHIN THE DYSTOPIAN IMAGINATION AND THE MARK: THE BEAST RULES THE WORLD

ANDREW WATKINS, FAYETTEVILLE STATE UNIVERSITY
MENTOR: DEAN SWINFORD

Abstract

Insularity of setting within the dystopia is a motif that is shared between works of dystopian fiction and works in the utopian imagination as a whole. Each dystopia or utopia is unique in how it creates and uses insularity of setting to facilitate the confinement of its subjected society to a prescribed set of social, political, religious, and economic practices. Timothy LaHaye and Jerry B. Jenkins use the framework of rapture fiction and the motif of insularity of setting to create an international dystopia in their novel The Mark: The Beast Rules the World. The global insularity conceived within The Mark is unique in light of the traditionally isolated insularity found in classic dystopian and utopian works. The study of insularity of setting in the works of More’s Utopia, Orwell’s 1984, and Bradbury’s Fahrenheit 451 can be applied to the insularity found in The Mark, contributing to a better understanding of how this motif is used in dystopian fiction and how it works to create a globally insular dystopia in The Mark.

Introduction

Timothy LaHaye and Jerry B. Jenkins’s novel The Mark: The Beast Rules the World is a dystopian work of rapture fiction from the greater Left Behind series of novels. Identifying the work as rapture fiction is important to understanding the context of the dystopian setting within the novel. Rapture fiction is a subgenre of apocalyptic fiction that Matthew Guest identifies as being “ultimately inspired by John Nelson Darby’s 19th century formulation of the end times, characterized by the rapture, tribulation, rise of the Antichrist, battle of Armageddon, and culminating in the triumphant second coming of Jesus in advance of the millennium” (1). Knowing that the setting of the novel takes place in a postrapture, premillennial apocalyptic world helps create a dystopian frame of mind for the reader. The reader quickly realizes that the novel’s setting is dystopian because of the rapture-fiction framework as well as the recognition of Satan indwelling in Nicolae Carpathia as the world’s all-powerful leader.
In this setting, a key dystopian motif that contributes to the novel’s framework and plot is insularity of setting. Insularity of setting within the utopian imagination can be traced back to Thomas More’s *Utopia* and can also be traced in the settings of classic dystopian novels such as George Orwell’s *1984* and Ray Bradbury’s *Fahrenheit 451*. Insularity of setting within the utopian imagination helps to confine the social, political, religious, and economic areas of a society, all of which that society’s citizens must partake of, or else face exclusion, expulsion, punishment, or even death. A unique aspect of the insularity found within LaHaye and Jenkins’s *The Mark* is that it is global, forcing all people around the world to either take a loyalty mark in order to live and be part of the global society instituted by Nicolae Carpathia, or face death from the society’s authorities. This insularity is not confined to a specific geographical location as found in the traditional utopian or dystopian imagination such as More’s island of Utopia, Orwell’s Oceania, or Bradbury’s futuristic city in *Fahrenheit 451*. It is found in the all-encompassing system of government enacted by Nicolae Carpathia and the Global Community. The only escape from the dystopian insularity in this postapocalyptic setting comes from the shared hope of those people who have become Christians during the tribulation and await the return of their savior, Jesus Christ. In this way, *The Mark* is also unique. The hope for escape from dystopia to a possible utopian future is not able to be brought about through any human action other than faith in Jesus Christ. This faith promises an eternal utopia in Heaven, which provides motivation to the novel’s protagonists, the Tribulation Force, in their pursuits of evangelizing others and ministering to fellow believers.

Through examination of scholarly criticism on *The Mark* specifically and the Left Behind generally, an evident lack of scholarship exists concerning the novel and series as works of dystopian fiction becomes clear. Some criticism exists concerning theological matters and eschatological beliefs displayed in the novel and by its authors, but this criticism does little to qualify the value of the novel and the series as works of literature within the dystopian imagination, even if that imagination is fueled by deeply held religious beliefs from the authors and primary readership alike. Much other criticism of the novel and series, however, can be applied to the study of the dystopian framework contained in *The Mark*. The novel’s background as a work of rapture fiction, the unified global economy in place in the novel, and the use of technology, both to help facilitate global insularity by the antagonist and to assist in evangelism by the protagonists, are all discussed in criticism of the novel and can all be applied to the study of insularity of setting in the novel. Through examination of the dystopian motif of insularity of setting within *The Mark* and its similarity to the insularity of setting found within other dystopian and utopian works, a clearer understanding of how the novel operates as a work of dystopian fiction within the subgenre of rapture fiction is found.
Criticism Overview of The Mark: The Beast Rules the World

*The Mark: The Beast Rules the World* offers a unique blend of apocalyptic and dystopian fiction through the lens of biblical interpretation related to end-time events. Scholarship surrounding the novel and its greater series, *Left Behind*, offers a wide range of topics. Some topics focus on the theological views of the series’s authors, but others focus on the value of the novels as works of literature that offer interesting perspectives on social, economic, cultural, and technological themes. Some of these themes can be applied directly to studying the utopian imagination, specifically the dystopia. Essays from Matthew Guest, Andrew Strombeck, Felicia Wu Song, and Gorman Beauchamp provide a good overview of scholarship about or related to *The Mark* that starts from a broad historical context and narrows to a more focused look at the use of technology, which can be applied to the study of insularity of setting found within the novel and dystopian/utopian literature.

An informative overview of the *Left Behind* series of novels and their context within apocalyptic dystopian literature is found in Matthew Guest’s essay “Keeping the End in Mind: *Left Behind*, the Apocalypse and the Evangelical Imagination.” Guest provides scholarship on the historical context of the *Left Behind* series as a whole within the subgenre of rapture fiction produced in Christian literature (1). Through giving a background on the context of rapture fiction, Guest provides insight into the evangelical imagination that has produced it.

Because the evangelical imagination is grounded in certain eschatological beliefs, although these differ between the numerous divisions and denominations of Christianity, one approach to rapture fiction, such as the *Left Behind* series, has been to view it in the context of instrumentalization of the novel. Guest defines instrumentalization as “a tendency to articulate . . . significance chiefly in terms of a means to a particular end, most frequently assumed to be some kind of ideological agenda” (3). Guest proceeds to detail how much of the scholarship about the *Left Behind* novels revolves around the idea of the novels’ instrumentalization. This approach, however, does not attempt to fully understand or appreciate the novels as works of literature, and it also assumes that readers are either advocates for instrumentalization from the genre or those opposed to such instrumentalization, leaving no room for a general readership. Guest backs up the reasoning behind this instrumentalization idea with details from the *Left Behind* series authors, Timothy LaHaye and Jerry Jenkins. Guest mentions their beliefs, positions within evangelical Christianity, and previous work. His essay, however, goes beyond looking at the perception of the novels themselves and details their position in imaginative literature. Guest cites anthropologist Vincent Crapanzano for his view of the connection of apocalyptic ideas to the utopian imagination: “The utopian imagination is—suddenly, powerfully, and briefly—inflamed by the immediate prospect of radical change, by visions of an apocalypse now” (8). In this sense, *The Mark: The Beast Rules the World*...
the World, from the Left Behind series, showcases the postapocalyptic view of dystopian literature within the framework of rapture fiction and the utopian imagination.

In his essay “Invest in Jesus: Neoliberalism and the Left Behind Novels,” Andrew Strombeck discusses the cultural and economic basis for both the novels themselves and the interconnected society within the novels. Strombeck argues that neoliberalism and capitalism contribute to the series’s mass popularity. Instead of being an evangelical series that is written from a culturally isolated perspective, Strombeck notes, “The books refuse to condemn multiculturalism and in fact emphasize the multicultural composition of their protagonist group. The overall effect of these choices is to emphasize fundamentalism’s continuity, not discontinuity, with postmodernity” (165). Strombeck shows that the Left Behind characters are a multicultural group that reflects the global economy and the goal of evangelicalism to reach a global audience. There are characters in the Tribulation Force from a variety of races and nationalities. The books were mass-marketed and sold in the same globalized economy. The Mark: The Beast Rules the World has a plot that centers on this idea of a globalized society and economy. In these ways, Strombeck’s essay is helpful in understanding how economic and cultural factors play a role in The Mark.

Despite the Left Behind series’s popularity and multicultural characteristics, Strombeck makes an interesting note from another scholar, Melani McCalister, who has written about the books. McCalister states that the books have been “all but invisible in liberal and intellectual circles” (161). Strombeck uses this lack of visibility to emphasize arguments showing the dichotomy between the series’ global reach and its perceived narrow focus. The series also shows an unlikely positive view of technology, according to Strombeck, in that its protagonists take advantage of the latest technology to reach others with their evangelical message. This view of technology in positive terms related to the protagonists within a dystopian setting is somewhat unique in the genre.

To further study this view of technology and how it plays out in modern culture and in The Mark, a look at Felicia Wu Song’s essay “Being Left Behind: The Discourse of Fear in Technological Change” (2003) is insightful. Song discusses how the view of technology has changed over time from the period of the Enlightenment and the Industrial Revolution to modern times. The 2003 publication date of Song’s essay should not distract from the application of the essay’s main points when applied to 2023. If anything, the progress made in this twenty-year span serves to reinforce Song’s ideas. Song details how technology was once viewed from a more positive perspective of universal enlightenment but in modern times has taken on the ability to induce fear, especially of not being able to keep up in a globalized world of capitalistic economies. Song states, “I want to argue that the impetus behind technological adoption is not only the pull of the Utopian dreams of Progress, as many observers
have noted, but also the push of fear—specifically the fear of being left behind” (28). Song ties this idea of being “left behind” to the novels from the Left Behind series by LaHaye and Jenkins. The motif of insularity of setting in The Mark can also be tied to this idea of being “left behind” in that the general population is being coerced to take the loyalty mark in order to avoid being left out of society or killed off.

Throughout the series, and especially apparent in The Mark, there is a tension between how technology serves positive and negative roles. The loyalty-mark system is one facilitated by technology, but at the same time, the protagonists use technology to communicate with each other and with unbelievers. Song details, however, the connections between technological progress and the idea of the rapture narrative of being left behind. In both cases, there is a fear of being left behind, whether that is being left behind figuratively when not upgrading to the latest technology in business or personal life or literally when the rapture occurs and takes millions out of the world as in LaHaye and Jenkins’s novels. Song cites Robert Nisbet’s History of the Idea of Progress when stating, “Over the last two centuries, the Enlightenment heritage of Progress has by no means been constant or static in American culture” (29). This shows that although technology has progressed constantly, the ideas of whether this progress is positive or negative have changed. Many view technology from the perspective of fear of being left behind. The Mark has an interesting blend of using current technologies in ways that support the protagonists but also echo the idea of the fear of being left behind, especially in the case of those who choose to follow Carpathia out of fear.

With a more in-depth focus on technology and dystopian literature, Gorman Beauchamp’s essay “Technology in the Dystopian Novel” gives insight into this topic that can also be applied toward analyzing how insularity of setting is established in The Mark. Written many years before the Left Behind series, Beauchamp’s essay reviews the theme of technology in many important works of dystopian literature. In his work, differing approaches to technology are shown. Primarily, two fields of applying technology in the dystopia are shown. One shows technology as a tool being used by those in control to help maintain their control of society, such as in Orwell’s 1984. The other shows technology becoming able to control itself and dictate society, as in E. M. Forster’s “The Machine Stops.” Beauchamp states, “The typical view of dystopists nevertheless holds technology to be an autonomous force that dictates the ideology of the future” (57). There are variations to these views within each dystopia, but technology is a key element in the enforcement of ideology and insularity in dystopian literature.

Although not directly about The Mark or its authors, LaHaye and Jenkins, Beauchamp’s essay is about a key element used to facilitate insularity within the novel: technology. Technology serves the antagonist, Carpathia, in achieving his goal of a one-world government, so in this way, technology is used in the novel as a tool.
Technology is also used to monitor people and to keep track of who is part of the system. In a unique twist on this dystopian use of technology, the protagonists also readily use technology for the advancement of their causes. In these ways, Beauchamp’s overview of technology in dystopian literature can serve as a means to compare and contrast its use as a key element of the dystopian imagination in *The Mark*.

After review and analysis of criticism related either directly or indirectly to *The Mark: The Beast Rules the World*, the historical context and key dystopian features related to the globalized society and technology of this novel can be seen. LaHaye and Jenkins have written a unique piece of literature that combines biblical interpretation, the evangelical imagination, apocalyptic visions, and dystopian themes. The criticism of the novel, however, is light on direct study of literary elements found within it. More specifically, the study of *The Mark* as a work of dystopian literature is one area of scholarship that is lacking. A focused look at the dystopian motif of insularity of setting can contribute to bringing this novel and the series as a whole into the discourse on the utopian/dystopian imagination. The novel’s unique usage of the dystopian motif of insularity of setting, and how the motif is established, deserves further study by tracing insularity from a historical context in the utopian/dystopian imagination and showing how it relates to the insularity found in *The Mark*.

**Tracing Insularity of Setting in Dystopian/Utopian Works of the Past**

A look at prominent dystopian and utopian texts from the past will help to show how insularity of setting plays a key role in the dystopian/utopian imagination. Two novels that display insularity of setting within the dystopia are George Orwell’s *1984* and Ray Bradbury’s *Fahrenheit 451*. In Orwell’s *1984*, insularity of setting is displayed within the dystopian society of Oceania with its highly controlling government, extreme surveillance, and lack of vision to the outside world. The protagonist, Winston Smith, lives in London, which is part of Airstrip One in Oceania. Winston lives in this dystopian setting, giving the reader a view from the inside. From the very beginning, the narrator details the mindset Winston and others have about being watched by the government. The narrator shows the reader a poster that says, “BIG BROTHER IS WATCHING YOU” (Orwell 2), and then gives the reader insight about one method of surveillance, the telescreen: “The telescreen received and transmitted simultaneously. Any sound that Winston made... would be picked up by it” (3). By doing this, Orwell makes the setting close in on the reader, showing the insularity in which citizens of Airstrip One exist. This insularity of being watched in and confined to a world that does not allow for free thought or movement, or even vision to the outside world, except that which Big Brother allows through propaganda, is continued throughout the novel. This idea of watched confinement and its effect on humans is
displayed in an essay by Harry Strub titled “The Theory of Panoptical Control: Bentham’s Panopticon and Orwell’s Nineteen Eighty-Four.” In this essay, Strub concludes that “the theory of panoptical control is based upon the behavioral principle that so long as environmental cues can sustain the belief that one’s behavior is being constantly monitored… avoidance behavior will be obtained” (52). By creating insularity of setting through the use of surveillance and control tactics, Big Brother is able to obtain the obedience of its citizens within the dystopia.

In Bradbury’s Fahrenheit 451, insularity of setting within the dystopia is displayed in many ways. The burning of books by the firemen shows that insight into any other knowledge outside of what the authoritative society allows is prohibited. The protagonist, Guy Montag, becomes aware of the insularity only when Clarisse McClellan begins asking him many questions. Clarisse asks, “Do you ever read any of the books you burn?” to which Montag responds laughingly, “That’s against the law!” (Bradbury 5). McClellan asks question after question, showing a thirst for learning, which is unique in the insular society. Montag’s wife, Mildred, personifies the insularity in the novel by being obsessed with her parlor walls and earbuds. She lives her life in a seemingly neutral/geared existence of being fed junk-food entertainment visually and auditorily. Her obsession with the virtual world of her parlor walls reflects a concern about modern television and media that Daphne Patai details in her essay “Ray Bradbury and the Assault on Free Thought” when she states, “We see into other people’s lives (real or imaginary) in a way that seems increasingly unavailable in actual life” (42). Mildred is almost incapable of even wanting to interact with the real world. Mildred sees her entertainment as her actual family. Right after she turns her husband in for harboring books, she leaves their house and says, “Poor family, poor family, oh everything gone, everything, everything gone now” (Bradbury 108). Mildred is so caught up in the insular setting that virtual “family” members have become real to her. These virtual family members can be described as simulacra, which Helena Nee discusses in her essay “Manifestations of the Hyperreal in a Postmodern World.” Nee states that Mildred’s “goal is to stay in her ‘play’ world forever. She refers to the characters as her relatives and these people on the screen are now becoming… simulacra” (13). By showing insularity of setting in the prohibition of books and Mildred’s virtual reality, Bradbury is able to create a dystopia that its citizens are able to see outside of only if they ask questions or attempt to engage with reality, which is almost impossible to find, as all books get burnt up and reality becomes full of simulacra.

Insularity of setting can be traced back in the utopian imagination to Thomas More’s Utopia. The island of Utopia was literally and figuratively insular. In Utopia, Raphael Hythloday begins his description of Utopia by stating that “the island of the Utopians is two hundred miles across” (38). This first part of Hythloday’s account is important in creating the insular setting for Utopia. It is an island nation that is sheltered from the outside world geographically. Utopia has its own political system,
social practices, religious practices, and economic system, all of which must be followed by citizens. These systems and practices create insularity within the society of Utopia. Unlike the common view that comes from within dystopian insularity, Hythloday views Utopia’s insularity as an outsider. He sees the nation as idealistic; however, some of Utopia’s insularity is less idealistic and comes from penalizing transgressors or those who do not accept its systems and practices. Hythloday notes that if the population grows too large, Utopia will send citizens to the mainland to occupy unused land. There, “if the natives will not join in living under their laws, the Utopians drive them out… and if they resist make war on them” (More 49). This shows that Utopia is not as tolerant as the reader might expect from an ideal place. The nation protects its insular system, even outside of the island, showing that the systems and practices are what truly create insularity in the setting.

For citizens on the island of Utopia, travel is limited and needs to be permitted by local authorities. As Hythloday details, “Anyone who takes upon himself to leave his district without permission… is treated with contempt… and severely punished. If he is bold enough to try it a second time, he is made a slave” (More 53). Movement and free will are limited in this insular setting. One difference between More’s Utopia and modern utopian works is the historical utopia’s ability to be isolated and hidden from the world. In “Varieties of Literary Utopias,” author Northrop Frye comments that “technology tends to unify the whole world. The concept of an isolated utopia… evaporates in the face of this fact” (216). The reader of modern utopian literature can see that in order to create insularity, an author must create insularity within a society’s systems, similarly to the island of Utopia’s system of laws, which isolates its citizens from the outside world. In “Varieties of Literary Utopias,” author Northrop Frye comments that “technology tends to unify the whole world. The concept of an isolated utopia… evaporates in the face of this fact” (216). The reader of modern utopian literature can see that in order to create insularity, an author must create insularity within a society’s systems, similarly to the island of Utopia’s system of laws, which isolates its citizens from the outside world. One must accept the insularity of Utopia and its systems and practices in order to go unpunished. In this way, the insularity of setting within More’s Utopia shows characteristics similar to those in dystopian literature.

The motif of insularity of setting within the dystopia is prevalent throughout the history of the utopian imagination. Even in More’s Utopia, insularity of setting can be seen as potentially dangerous when in governmental control. Insularity can mean geographic and social protection in Utopia, but it can also mean limited freedom and isolation from ideas outside of the insular society. This idea of governmental control of insularity is taken to extremes in Orwell’s 1984 and Bradbury’s Fahrenheit 451. The dystopia begins where the utopian idea of insularity is magnified to a point where free will and freedom to think for oneself are diminished. In 1984, Big Brother uses surveillance and severe punishment to gain control of Oceania’s citizens, increasing the insularity in that society. In Fahrenheit 451, Captain Beatty and the government in control of Guy Montag’s society use book burning to suppress free thinking. Technology, in the form of parlor walls and earbuds, is also used to create insularity in Bradbury’s novel by detaching humans from reality, as personified in Mildred Montag, thus creating an insular society that does not want to face reality. The novel The Mark: The Beast Rules the World uses this same element of insularity within the postrapture
world of the Left Behind series. Insularity on a global scale is created as the world’s leader, Satan in Nicolae Carpathia’s body, establishes and enacts a loyalty mark that individuals must take in order to be part of the Global Community society. In The Mark, the majority of citizens are more than willing to be obedient to Satan’s system, partly because of the threat on their physical bodies and partly because of their infatuation with Carpathia’s power and persuasion. This obedience to governmental control is similar to that seen in the citizens of Utopia, Oceania, and the city of Fahrenheit 451. Tracing insularity of setting within the utopian imagination and applying it to The Mark makes apparent that although insularity can be tied to physical and geographical location, it is more clearly seen in the societal insularity of a suppressive government. The Global Community of Nicolae Carpathia shows that insularity can be on a global scale within the dystopian imagination.

Insularity of Setting in The Mark: The Beast Rules the World

In The Mark: The Beast Rules the World, by Timothy LaHaye and Jerry B. Jenkins, the dystopian motif of insularity of setting is envisioned on a global scale, connecting humans via a singular economic, religious, social, and political system that has zero tolerance for nonadherence. The antagonist, Nicolae Carpathia, embodies the global system that leads to creating the insular society and setting of the novel. The Global Community is the singular governmental system of the world, and Carpathia is the head of it. Carpathia is also the object of worship of the one-world religion, Carpathianism. Loyalty to the man, Carpathia, and to his system of government is preeminent. Part of this system, and the driving narrative of the novel, is the requirement for all people to receive a “mark of loyalty” (LaHaye and Jenkins 84), that “consist[s] of the name of His Excellency or the prescribed number” (85). Carpathia describes this process and the punishment if it is not accepted as “certain unfortunately necessary controls on the citizenry” (83). The reader and protagonists are aware that Carpathia is indwelt by Satan, but the majority of global citizens view Carpathia as an infallible god. The citizens of the world who willingly accept the mark, or plan to accept as soon as possible, are blind to any other viewpoint than that which Carpathia puts forth, reinforcing insularity in the global society.

The novel begins to quickly establish insularity of setting in chapter one by showing how Carpathia himself consumes the attention of all people. When David Hassid, one of the protagonists in the Tribulation Force, is introduced to the reader, he is being confronted with images of Nicolae Carpathia in the city of New Babylon: “He could barely turn his eyes from the gigantic screens. . . . The image of the indefatigable Nicolae Carpathia, freshly risen from three days dead, filled the screen and crackled with energy” (LaHaye and Jenkins 1). The role of technology is used to help distribute the insularity of setting in a way that is reminiscent of Orwell’s telescreens and Big Brother posters in the beginning of 1984. David, in his search for Annie, is confronted
not only with this broadcast of Carpathia but also with numerous citizens excitedly scrambling to see Carpathia in person. Carpathia has developed his devoted following through the mass communication of his image and resurrection. This development of a following via mass communication and propaganda is like what Anita Pisch describes concerning the ability of such personality cults to develop in modern times: “Modern personality cults are possible due to the capability to disseminate images of the leader and his achievements over wide distances and to saturate public space” (50). The personality cult of Carpathia helps create insularity by consuming the affections of the masses. David Hassid does not join in the Carpathia fever, beginning a pattern of protagonists in The Mark who live in the dystopia but avoid becoming part of it, giving the reader a glimpse from inside the insular dystopian setting.

Hassid is located in New Babylon, a Middle Eastern city, but the insularity of Carpathia’s system is not limited geographically to this area. The novel switches back and forth between Hassid in the Middle East and the core group of the Tribulation Force, who are located in Chicago. The Tribulation Force is based in an abandoned office building in Chicago, where they organize missions and communicate via computers, phones, and radios to other Christians in order to help each other survive within the dystopian setting and impede the efforts of Carpathia to establish and enforce his Global Community system. David Hassid provides inside information and access to the Global Community’s computer systems for the Tribulation Force. In this way, the insularity of setting is stretched across the globe. All people are subject to the same repressed society and suppressive government, or what could be described as a dystopian “global monoculture” (Baccolini and Moylan 8). The Tribulation Force is subject to being spied on and found out by the same Global Community forces as those active in New Babylon. Rayford Steele tells Albie, “Never know what satellite imaging shows…. We could be sleeping soundly while GC Security and Intelligence forces snap our pictures from the stratosphere” (LaHaye and Jenkins 8). Once again, insularity of setting is facilitated through technology, creating a dystopia that the protagonists must secretly work within in order to help other Christians who do not want to become part of Carpathia’s system.

As the novel continues, insularity of setting continues to be shown as Rayford Steele and Albie work as undercover Global Community officers in their attempt to rescue Hattie Durham, a nonbeliever whom the Tribulation Force has been trying to lead to becoming a Christian. Hassid assists in this by infiltrating the computer systems of the Global Community, giving the men credentials and permissions to operate within the dystopia: “David often overrode other GC systems to send such directives in a way that they could not be traced back to him” (LaHaye and Jenkins 35). The protagonists cannot move freely within the society without either hiding from the Global Community or appearing as one of the GC. The permissions needed to move about closely resemble the permissions needed to move between regions on More’s island of Utopia. Christian protagonists in The Mark are all constantly at risk of being
found out, which would lead not only to their own death or punishment but possibly also to the demise of others in the protagonist group.

Nevertheless, the Tribulation Force members risk their lives in hopes of evangelizing others who have not committed to the Global Community, showing a hope in the utopian future they believe is coming at the second coming of Christ. This utopian hope within the insularity of a dystopia is a foundational principle of utopianism as described by Darko Suvin: “Utopianism is an orientation toward a horizon of radically better forms of relationships among people” (187). LaHaye and Jenkins have taken the basic principle of utopianism and displayed it within the insular dystopian setting of *The Mark*. The protagonists are oriented toward a utopian horizon even while being in a dystopian world ruled by Satan.

The motif of insularity of setting builds to the point where Carpathia and his constituents come up with a plan that further enforces insularity through a system in which each person must take a loyalty mark. The mark of loyalty symbolizes everything Carpathia stands for and grants access to his insular society. This access is facilitated by the technology embedded in a biochip received in the mark, which Carpathia’s loyal assistant, Viv Ivins, details: “We have settled on the technology…. The miniature biochip with the suffix numbers embedded in it can be inserted as painlessly as a vaccination” (LaHaye and Jenkins 86). And without the biochip, people “will not be allowed to buy or sell” (85). The entire world is subject to this loyalty system, and the numbers embedded in each mark are connected to “a listing of ten world regions… identifying the home regions of every citizen. The subsequent numbers… will further identify the person to the point where every one shall be unique” (85). This idea of everyone being unique contrasts sharply with the fact that all citizens will be sharing the same exact system of values, beliefs, economics, and political views.

As seen in other dystopian settings, citizens who do not follow the exact prescription for how to live life in such a suppressive society are punished. This loyalty mark system in *The Mark* sets up an insular society that beheads those who openly refuse it via “loyalty enforcement facilitators” (LaHaye and Jenkins 170), which are guillotines, and grants no access to the economic system for those who hide from the Global Community. The Tribulation Force and fellow Christians become further isolated within the dystopia and must hide in order to avoid having to openly refuse the mark and face death. This punishment process facilitates the creation of insularity of setting with medieval technology, and this technology is displayed openly to the public, as can be seen when Laslos discusses the parading of guillotines: “The GC is carting them through town in open trucks so the people can see the consequences of thinking for themselves” (263–264). The suppression of free thought in *The Mark* further strengthens insularity of setting in a way similar to that of Bradbury’s firemen, who burn books, suppressing access to free thought. At another point in *The Mark*,

332
Ming Toy pleads with David to help free her brother, Chang Wong, and confesses she is “thinking with [her] heart” (279), to which David responds, “Nothing wrong with that . . . until we quit thinking at all and make things worse” (279). Once again, free thought stands in contrast to the ideology of an insular dystopian society.

The harshest realities of creating insularity of setting are seen in *The Mark* when the beheading process begins at prison facilities. Buck Williams and Albie are undercover, trying to free Christians in a prison beginning the loyalty enforcement process when they witness the beheadings of several who refuse the mark. Insularity of setting becomes even more real through this gross display of disloyalty punishment. As Maria Varsam states in her essay “Concrete Dystopia: Slavery and Its Others,” “In dystopian texts, public as well as privately enforced violence and its threat ensure the obedience of the less powerful” (212). The prisoners are exposed to the immediate threat of capital punishment for refusing the loyalty mark, which leads to compliance from most. Mrs. Miklos and the first women to refuse the mark and accept death exemplify Christian martyrs of times past as the narrator states, “They would join the great blood-washed who literally made their bodies living sacrifices” (LaHaye and Jenkins 302). Mrs. Miklos is beaten horrifically before being led to the guillotine. As she is going, Buck passes along a message from her husband, to which she says, “Tell Laslos thank you for leading me to Jesus. I see him. I see him. I see my Savior and can’t wait to be with him!” (303). Mrs. Miklos, even while being beaten and led to her death, has a hope in a utopian future that her current imprisoned, beaten, and death-bound dystopian existence cannot touch. As the narrator notes at the end of this scene of beheadings, Buck “stood in the cool grass, convulsing now in dry heaves, [and] covered his ears in a vain attempt to muffle the thuds and cheers, thuds and cheers” (305). LaHaye and Jenkins show the ugliness of the dystopian society that cheers at the death of others. Through this loyalty enforcement, insularity of setting is shown to ultimately lead to a choice between taking part in the dystopian society or accepting death or dismissal from the society.

**Conclusion**

The utopian/dystopian imagination has a long literary history filled with variations of general motifs and common elements that adapt to the times, desires, and visions of the authors creating it. Insularity of setting is one motif that shares both classic utopian literary usage, as traced in More’s *Utopia*, and dystopian usage, as traced in Bradbury’s *Fahrenheit 451* and Orwell’s *1984*. By the tracing of its usage and how it facilitates the creation of both utopian and dystopian worlds, insularity of setting is shown as a key literary element within the entire utopian imagination. Typically, insularity of setting is used in the more negative sense, especially in the dystopian imagination. Although utopias and dystopias are polar opposites, or extremes, it is clear that insularity of setting shows itself in both cases.
In the dystopian world, through tracing insularity of setting in the works of Orwell, Bradbury, and now LaHaye and Jenkins, it is apparent that this motif is often facilitated by technology in some way, acting as a unifying and controlling force. Orwell’s dystopia in *1984* is filled with references to technology being used as surveillance to help control the citizenry. In Bradbury’s dystopia in *Fahrenheit 451*, technology is used to control the citizenry through mind-numbing information and entertainment, as displayed by the parlor walls. In LaHaye and Jenkins’s *The Mark*, technology is used to unite all people through a loyalty mark system, facilitated by an injectable biochip and a database that creates a one-world community. The creation of insularity of setting, however, is most dependent upon the enforcement of a strict system of social, economic, political, and religious practices. This system is what actually establishes insularity, and it can be seen in both dystopian and utopian instances.

By studying insularity of setting in its historical context within dystopian and utopian works, we can better understand it as an extreme, confining system that must be adhered to by citizens within the utopia or dystopia and that, if they do not adhered to it, the citizens are excluded, expelled, punished, or even killed. *The Mark* displays this type of insularity brilliantly on a global scale using the framework of rapture fiction. Insularity can be thought of as coming from isolation from the rest of the world, and this is true in a sense, but within *The Mark*, insularity of setting can be seen better as a system that confines its citizens to a globally unified set of social, political, economic, and religious beliefs. In this way, the entire globe is isolated in the Global Community and Carpathianism. Through the tracing and studying of the dystopian motif of insularity of setting within *The Mark*, the novel can be better understood as a work of dystopian literature coming from the subgenre of rapture fiction. *The Mark: The Beast Rules the World* can then be viewed and appreciated as a work of dystopian literature that draws from the same fountain as other utopian and dystopian works, but in a way that is unique to the vision, desire, and times of its authors.
Works Cited


