

The

PIONEER RESEARCH

Journal

An International Collection of
Undergraduate-Level Research

Volume 8

2021

Pioneer[®]
academics

Copyright 2022, Pioneer Academics, PBC, all rights reserved

Publisher

Hong Kong Digital Publishing Ltd.
1501, Grand Millennium Plaza, 181 Queen's Road, Hong
Kong

ISBN 978-988-77066-7-0

Pioneer Academics, PBC
Email: info@pioneeracademics.com
Website: www.pioneeracademics.com

The Pioneer Research Journal is published annually by Pioneer Academics, PBC. The 2021 issue of *The Pioneer Research Journal* is Volume 8.

Copyright ©2022, by Pioneer Academics, PBC, 30 S. 15th Street, 15th Floor, Philadelphia, PA 19102, USA. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Contents

<i>Foreword</i>	<i>ix</i>
<i>Selection Process</i>	<i>xi</i>
Being’s Chain, God’s Death, and Atlas’s Shrug: How Meaning Is Blurred and Deferred in the Consideration of Godlessness and Freedom (Philosophy)	1
Author: Anna Kassim School: Croydon High School – Croydon, London, United Kingdom Pioneer Research Concentration: The Examined Life	
Race in the Eugenics Movement and the Progressive Era, 1910-1939 (Science, Technology, and Society)	22
Author: Ariane Vidal School: San Francisco University High School – San Francisco, California, United States Pioneer Research Concentration: Modern Technology: The Future, Present and Past	
Film’s Portrayal of Mental Illness and Violence: Schizophrenia DID, and PTSD (Media Studies)	35
Author: Brian Nguyen School: Northside Preparatory High School – Chicago, Illinois, United States Pioneer Research Concentration: Media and Society	
Making Decisions in a Crisis: The Ethics of Scarce Resource Distribution Policies During COVID-19 (Philosophy)	48
Author: Caroline Greenstone School: Greenhill School – Austin, Texas, United States Pioneer Research Concentration: Bioethics	
Myanmar Risk Analysis: McDonald’s Case Study (International Relations)	61
Author: Catherine He School: Ridley College – St. Catherines, Ontario, Canada Pioneer Research Concentration: Global Risk and Prediction	

- A Contemporary Perspective on the Health of the Great Barrier Reef and the Wellbeing of the Torres Strait Islanders** (Environmental Studies).....75
Author: Chong (Phelps) Tin
School: Eton College – Windsor, Berkshire, United Kingdom
Pioneer Research Concentration: Global Indigenous Health
- Using Networks and Supervised Machine Learning to Analyze and Predict the Results of Team Sports Games** (Computer Science).....105
Author: David Li
School: William P. Clements High School – Sugar Land, Texas, United States
Pioneer Research Concentration: Graph Analytics: Algorithms and Machine Learning
- Imagining a Carbon-Neutral Future Through Standard Oil’s Last Four Sisters Standing: Beyond Petroleum, Chevron, ExxonMobil, and Shell** (Environmental Studies).....125
Author: Dorothy Chan
School: Brooklyn Technical High School – Brooklyn, New York, United States
Pioneer Research Concentration: Waste or Not Waste, That’s the Question – Solving Major Environmental Problems
- Putting Constraints on the Dark Matter Density at the Center of our Galaxy from the Most Recent Observations of GRAVITY** (Physics/Astronomy).....146
Author: Efe Uğurlu
School: Robert College – Istanbul, Turkey
Pioneer Research Concentration: Relativity, Cosmology, and the Contents of the Universe
- The Consequences of Urbanization on Access to Housing and Education in Turkey: Struggles and Coping Strategies of Syrian Refugees** (Sociology/International Relations).....162
Author: Ezgi Erzincan
School: Lycée Saint Joseph – Istanbul, Turkey
Pioneer Research Concentration: Globalization and International Migration

- A Proposal to Facilitate Endogenous, Physiological Adrenal Steroidogenesis in 21-OHD CAH Patients via Recurrent Enzyme Replacement Therapy (Biology).....189**
Author: Isaac Lowry
School: Granville High School – Granville, Ohio, United States
Pioneer Research Concentration: Cell Biology
- Along a Post-Colonial Spectrum: Considering Theory through an Analysis of Jordanian and Lebanese Identities (International Relations).....229**
Author: Isabella Lashley
School: The Chapin School – New York, New York, United States
Pioneer Research Concentration: Future Global Orders
- Effectiveness of Virtual Reality in Restoring Arm Function Following Stroke (Biology/Neuroscience).....246**
Author: Kenedy Quandt
School: Sierra Canyon High School – Chatsworth, California, United States
Pioneer Research Concentration: Treatment of Arm Dysfunction After a Stroke
- Opening Our Hearts: An Analysis of How Noel Pearson Reframes the Debate through Irony and Identification (Political Science/History).....269**
Author: Laura Zhang
School: Ravenswood School for Girls – Sydney, NSW, Australia
Pioneer Research Concentration: Rhetoric and Social Movements
- Trans Selfhood, (Mis)recognition, and Post-Identity Resistance (Philosophy).....283**
Author: Moxuan Xiong
School: Rutgers Preparatory School – Somerset, New Jersey, United States
Pioneer Research Concentration: The Examined Life
- An Examination of Descartes' Doubt and the Foundations of Certain Knowledge (Philosophy).....297**
Author: Nathaniel Chapman
School: St. Paul's School – Concord, New Hampshire, United States
Pioneer Research Concentration: Descartes' Meditations on First Philosophy

- The Relationship Between Religious Practice and Depression: Intrinsic Religiousness as Moderator (Psychology)**.....305
Author: Qianyu Liang
School: Shenzhen College of International Education – Shenzhen, China
Pioneer Research Concentration: Stress and Health
- Structure-Based Mixed Cannabinoid Design Presents New Opportunities for Fibrosis Treatment: a Computational Study (Chemistry/Biology)**.....324
Author: Sangwon Shin
School: SMIC Private School – Shanghai, China
Pioneer Research Concentration: Using AutoDock Vina to Design Next-Generation Small Molecule Therapeutics
- Exploring the Adverse Health Effects Associated with Atrazine Exposure and Whether More Substantial Restrictions Are Required in the United States (Environmental Studies)**.....352
Author: Simratpal Deol
School: Ithaca Senior High School – Ithaca, New York, United States
Pioneer Research Concentration: Environmental Challenges: Environmental Choices
- Mutability and Function of SARS-CoV-2 ORF9b (Biology)**.....373
Author: Xiaoyi (Charlotte) Shi
School: Shanghai Foreign Language School Affiliated to SISU – Shanghai, China
Pioneer Research Concentration: Pandemic: The New Coronavirus
- Expanding Intergovernmental Water Markets Between Colorado River Basin Tribal Countries and Western States as a Response to Socioeconomic Drought (Environmental Studies)**.....401
Author: Xinran (Cici) Xu
School: Saratoga High School – Saratoga, California, United States
Pioneer Research Concentration: Environmental Justice and Resource Conflicts
- Comparing the Response Rate, Remission Rate, and Effective Age Range of High-Frequency Left Transcranial Magnetic Stimulation (HFL rTMS) and Low-Frequency Right Transcranial Magnetic Stimulation (LFR rTMS) in Treatment of Major Depressive Disorder (MDD) (Psychology)**.....419
Author: Xintong Chen
School: York School – Monterey, California, United States
Pioneer Research Concentration: Clinical Psychology

#ChallengeAccepted for Women in Turkey: Trending Instagram Participation or Collective Feminist Activism? (Sociology/Media Studies).....	433
Author: Yong Jie (Justina) Chou School: St. Paul's Co-educational College – Hong Kong, China Pioneer Research Concentration: Social Media and Human Progress	
The Role of Iconoclasm in the History of Byzantine Art (Art History).....	454
Author: Zehra Gundogdu School: Bahçeşehir High School for Science and Technology – Istanbul, Turkey Pioneer Research Concentration: The Medieval Mediterranean: Confluence of Cultures	
EDU Valuation: A New Era for Private Education or the Beginning of the End? (Business).....	477
Author: Zihao Gao School: McCallie School – Chattanooga, Tennessee, United States Pioneer Research Concentration: Financial Accounting & Analysis, Capital Markets, Entrepreneurship, and Decision Making	
Migration's Effect on the Sending Country: The Case of Poland (Economics).....	499
Author: Ziyu Zhou School: BASIS International School Park Lane Harbour – Huizhou, China Pioneer Research Concentration: International Economics	

Foreword

Since its inception, the *Pioneer Research Journal* has featured the work of young scholars driven by the desire to engage with their world through the pursuit of knowledge. The scholars whose work is featured in Volume 8 continue the tradition of academic excellence established by those who came before them as they add their voices to ongoing debates of vital importance as well as new debates arising from the COVID-19 global pandemic. Whether confronting a current social or political crisis or offering a revitalized perspective on an age-old philosophical question, each of the 26 research papers in this volume makes an important contribution to our collective knowledge and understanding of the world in which we live.

The papers contained in this multidisciplinary volume are the culmination of intensive research conducted by the global community of high school scholars who participated in the Pioneer Research Program. In 2021, 1,178 scholars from 48 countries and regions conducted research through Pioneer Academics, selected from an international pool of 4,256 applicants. Those admitted to the program participated in a faculty-led, international cohort before working one-on-one with leading U.S. professors to conduct original undergraduate-level research in their area of interest. After a rigorous nomination and double-blind review process, 26 papers were selected for publication in this edition of the *Pioneer Research Journal*. The authors represented herein are from schools in Australia, Canada, China, Turkey, the United Kingdom, and the United States.

The partnership between Pioneer Academics and Oberlin College is built on a shared commitment to superlative academic quality and diversity in education. By conducting the program entirely online, Pioneer Academics has torn down barriers and made undergraduate-level intellectual exploration available to promising young scholars in virtually every corner of the world, while its collaboration with Oberlin ensures continued excellence in faculty pedagogy and student learning outcomes. To ensure this opportunity is available to as many students as possible, Pioneer provided US \$1.16 million in need-based scholarships, again meeting 100% of demonstrated financial need where need could be assessed, thereby fulfilling its mission of removing obstacles to educational engagement for the most deserving underserved students.

The papers in this year's journal reflect the academic standards that define the Pioneer Research Program. Moreover, the willingness of these young scholars to build upon and to challenge the ideas of the thinkers with whose work they so skillfully engage demonstrates the heights that can be reached when learners' interest and ambition are harnessed and coupled with expert guidance.

We are proud to share this work widely. The *Pioneer Research Journal* is available in print and online at www.pioneeracademics.com and is distributed to select colleges, universities, and libraries worldwide.

I am pleased to share it with you and hope you find it to be inspiring and enlightening.

David G. Kamitsuka, Ph.D.

Dean of the College of Arts and Sciences
Oberlin College

Editor, *The Pioneer Research Journal*, Volume 8

Selection Process

To be nominated for publication in *The Pioneer Research Journal*, Pioneer scholar-authors first had to earn a letter grade of A- or higher from their professor mentor. Each professor mentor was invited to nominate one or two papers (depending on the number of scholars they mentored) that met the A/A- grading threshold and that, in their estimation, represented the highest caliber of scholarship from among their mentees.

Following nomination, every paper was reviewed anonymously by a member of our 44-member committee of reviewers, each of whom is a professor with expertise in the subject area of the papers they reviewed.

Contributing Readers scored the papers on a scale of 0 – 6 for four criteria: Engagement with Scholarship, Evidence & Analysis, Writing & Organization, and Scholarly Contribution. These four scores were then tallied, with a possible maximum score of 24. Papers with sub-scores of 1 in any category were disqualified from publication.

Once the above thresholds were met, the highest-scoring papers were provisionally selected. Upon notification, authors of provisionally selected papers received instructions from their paper's reviewer specifying the revisions required for meeting the standards of an undergraduate research journal. Every paper then underwent one more rigorous review by Pioneer's Writing Center tutors for editorial concerns and the authors were asked to make final revisions based on that review before their papers were finally published.

When all criteria were applied, only 2.3% of the papers generated in the 2021 Spring-through-Summer and Summer-Only terms were selected for publication in Volume 8 of *The Pioneer Research Journal*.

Being's Chain, God's Death, and Atlas's Shrug: How Meaning Is Blurred and Deferred in the Consideration of Godlessness and Freedom

Anna Kassim

Author Background: *Anna Kassim grew up in England and currently attends Croydon High School in Croydon, London in the United Kingdom. Her Pioneer research concentration was in the field of philosophy and titled "The Examined Life."*

Abstract

Language is an important tool for expressing thought, and the history of ideas stands as a testament to this. However, our use of language, and our perceived world, form a feedback loop where numerous, and sometimes unnoted, meanings are blurred and deferred while undergoing constant change between the literal and figurative. This phenomenon is especially prominent in the intellectual evolution that links God, godlessness and freedom—paradigmatic case studies. This paper takes this intellectual evolution and offers a potential narrative that posits and evaluates an epistemological and linguistic explanation for meaning's undetermined fluidity. Through reference to literature, including plays, novels, and other forms of academic expression, I establish an intellectual genealogy supporting my claims. I conclude this paper with an appeal to the expansion of definitions and concepts as well as a critical review of the hesitation and unease that marks this process.

1. Introduction

In language, there are different kinds of meanings, and these evoke different responses. In writing, the boundaries between literal and figurative meaning, what the writer intends, and what the reader interprets can be blurred. Thinkers across numerous intellectual disciplines—philosophy, politics, literature, art—have blurred the literal and figurative in an attempt to convey their ideas and make sense of what they hold to be corresponding reality. Despite this, the linguistic and epistemological grounds upon which such tendencies are based remain a consideration rarely touched upon or explored. In this paper, the underlying motivations that lead to these tendencies will be unveiled as a matter of both linguistic and epistemological concern and uncertainty. This will be done through an analysis of historical discussions—in the form of philosophy, plays, and books—concerning God, who will later come to be redefined as 'God.' His Chain of Being and death, as well as godlessness and freedom. The intellectual

mediums mentioned were chosen to demonstrate the simultaneous centrality of language to the way the ideas are conveyed and the tendency for language to defer and blur between the limits of perceived literality and figurativeness. The broad topics of God, godlessness, and freedom extend across all domains of social, cultural, and academic expression; therefore, it is not feasible to cover all these topics in one research paper. Nevertheless, this paper seeks to provide a plausible explanation and genealogy for the conceptual division to which language and the history of ideas have given rise.

Intellectual discussion of God(s) tends to focus on the possibility of His or their existence(s) and ponder what may be the attributes, role, and purpose of God, as well as how such an existence can be defended against a backdrop of alternative, competing metaphysical explanations. It is less common that discussion should focus solely on the use of language to define our literal and figurative perceptions of God, outside of discussions concerning religious language. Nietzsche's philosophy is one of the most famous examples of a discussion in which the focus was not solely on the possibility of God's existence or our ontological relation to Him. Whilst the scholarly interpretation of his work is diverse, this paper finds its groundings in an interpretation of Nietzsche's work as an intellectual pondering God as the sole source of morality, objectivity, and the consequences for a godless man, with the hope of further extending the discussion to include more linguistic consideration. Thus, the aim of this paper is to explore the intersectional sub-clause of language, meaning, and uncertainty within the greater topics of God, godlessness, and freedom. Intellectual positioning—both historical and intellectual—also weighs heavily on the research and conclusion of this paper and shall be reviewed in a later section.

The outline of this paper consists of five main sections of a selectively pragmatic and focalized genealogy of the history of ideas and their linguistic components. Before exploring the different subtopics in this paper, the boundaries of the definitions of literality and figurativeness will be explored to demonstrate what will be used as the basic definitions for the terms in this paper. The first section of this essay is an exploration of how one may define God, both in the literal and figurative sense. Its focus is on how defining God with a holistically figurative definition is more reflective of the term's use than an assumed religious definition. Sophocles' *Oedipus Rex* will be analyzed to demonstrate how we define God in both the literal and figurative. A grammatical distinction will then be made in the remainder of the essay to distinguish between this new figurative definition, 'God', and the previously assumed, traditionally religious use of the term, God. The next section concerns the death of God, considering its intellectual and linguistic nature as well as its relation to the conceptualization of 'God' and freedom. Having introduced God's death, the third section will focus on the definition of freedom and the intellectual tangles that have posited its literal aspiration and figurative ideal. Ayn Rand's *Atlas Shrugged* will be analyzed as a literary paradigm of the blurring of literal and figurative meaning and intention. This section will then question the desirability of freedom and its relation to the apprehensive acceptance of the death of God. This paper also explores Arthur Miller's *The Crucible* in order to draw parallels between historical systems of control and order. The fourth section will seek to explain the setting in which the discussed use of language is able to take on such fluid forms, as well as review the distinction made between definitions of the literal and figurative.

The fifth section of this paper shall evaluate potential challenges and objections to claims made in this essay before offering a final review of this paper's aim, selected content, and historical standpoint—before a final conclusion is drawn from its findings.

2. Literal versus Figurative

There are two ways in which the meaning of language shall be assessed: Contextually and Intentionally. The first way of discerning between the literal and figurative would be to assess whether language, put together in sentences, for example, may be said to reasonably correspond with perceivable reality. This way of distinguishing between the literal and the figurative may also be expressed using Colin Turbayne's example in which he notes that grammar is not itself a definite marker of this boundary, by considering this:

Table 1. *Lack of grammatical distinction in metaphors (Turbayne, 1962)*

The timber-wolf is a wolf	Man is a wolf
That frying-pan is a utensil	Mussolini is a utensil
Ordinary language	Visual language
Muscle fatigue	Metal fatigue

Yet, I am inclined to say that only the second members involve metaphors. (Turbayne 14).

Indeed, this inclination reflects how perceivable reality may act as a contextual barrier and signifier between the literal and figurative, including the metaphorical. Nevertheless, intentions still blur our perception of the boundary between the literal and figurative, especially in less obvious examples. Turbayne argues that when "my words are not to be taken literally but only metaphorically... and I implicitly ask my audience to do the same", a metaphor is created (Turbayne 14). This argument of what a metaphor is evades the risk that we may be fooled into thinking that the figurative is literal or vice versa. There may also be those who hold all language to be necessarily figurative, making Turbayne's and other attempts to distinguish between the two dimensions of meaning futile; this viewpoint shall be revisited in a later section. The latter way of discerning between meaning—intentionally—shall be explored and defined using literary examples. The consideration of God exemplifies how intellectual evolution and changing epistemology and metaphysics can make this discernment between the literal and figurative—both as a linguistic and judgement-based task—confused, uncertain, and possibly unachievable.

3. God or 'God'

What encompasses a definition is debatable. This is true especially if a word spans both the literal and figurative dimensions of meaning. No historical figure looms larger in this consideration or straddles the two dimensions more than God, or what will come to be referred to as 'God', especially considering His incomplete academic and cultural life cycle.

Indeed, for much of philosophical and social thought preceding what may be referred to as modernity, God(s) were those which purposefully, causally, unequivocally, and literally determined the flow of the tides. Gods also ordered beings in accordance with their divinely ordained Chain, which correlates to the ontological 'hierarchy' of creation. Hubert Dreyfus and Sean Dorrance Kelly insightfully point out in their book concerning our 'post-God' nihilistic period that "... the idea that everything had its proper place in God's divine plan was not itself a belief one could accept or reject. It was an entire way of life" (Kelly 14). God was omnipresent but his place in life was unquestioned. Even further, they argue that "the way of life of a culture is not an explicit set of beliefs held by the people living in it. It is much deeper than that. A person brought up in a culture learns its way of life the way he learns to speak in the language and with the accent of his family and peers" (Kelly 14). God in this sense may be understood as a language which through cultural embedding is inescapable, where every language move one makes is accented by 'God-speak', literal yet inconspicuous. This may suggest that despite the belief in God being literal, its inconspicuousness and standardization already counter claims that the literal is any more consciously accessible or recognizable than the figurative.

Nonetheless, the figurative enters divine conversation when thinkers—including playwrights, poets, politicians, philosophers, and artists—try to make sense of this Chain of being and what may literally happen to those who find their place along that Chain. A paradigmatic example of how cultural embedding, God, and man were reunited and cosmically divided is Sophocles' *Oedipus Rex*, a meta-textual showcase. The argument of the play is that Protagoras' maxim, which Oedipus preempted, "Of all things man is the measure" ("Protagoras (Stanford Encyclopedia of Philosophy)"), comes crashing down under the weight of the gods' supremacy. The moral of the play was to obey the gods, accept their ordained sequences of events, and humble the human tendency to self-proclaim. The protagonist of the play and King of Thebes, Oedipus, opens the first episode with a foreshadowing warning to the chorus: "You pray to the gods? Let me grant your prayers" (Sophocles et al., line 245).

However, it is the climactic Exodus of the play which provides the best example of how the figurative informs the literal. Following Jocasta's suicide and Oedipus self-blinding, the Chorus, who by convention in their *stasimons* break the fourth wall, providing their own meta-textual analysis of the myth of Oedipus as a figurative example of what may literally happen to the audience if they transgress against the gods. Echoing the Athenian statesman Solon, the last line of the play—sung by the Chorus—sums up their analysis: "Now as we keep our watch and wait the final day, count no man happy till he dies, free of pain at last" (Sophocles et al., lines 1683-1684). It is indeed the gods who determine happiness, pain,

life, death, and everything in between, and Oedipus shows us what happens when one rages against fatalism. This would have struck a religious and didactic chord with the audience, suggesting that, contrary to presumption and the case of embedded God-speak, the figurative may more glaringly and conspicuously inform one about supposed reality than the literal, a revelatory possibility. It is, therefore, clear from this Ancient Greek tragedy that myth in this early incarnation was used in much the same way as the Bible was in later centuries: to govern, control, and provide genealogy, explanation, reasoning, order, and power. Although the parallel between monotheistic Abrahamic faiths and (predominately) polytheistic Greece is not perfect, the sense in which both involve divine authority, control, and myth allows for their philosophical relation. Indeed, a single Greek play may not have been held to the same hermeneutically singular standard the Bible was, and for some is, but its literary and linguistic tools warrant critical comparison. Similarly, the intellectual foundations of Abrahamic thought find certain roots in post-Socratic ancient thought.

The purpose of examining this play is to demonstrate how the figurative has been used to explain and make sense of what is believed to be literal and real, as well as show how the boundaries between the figurative and literal may become blurred. This meaning deferral has occurred to the extent that what is purely demonstrative or explanatory, what is truly meant by our use of language, and what is literally believed have become blurred. For example, the Greeks had many epithets for their gods, and many myths to accompany them to demonstrate the division of attributes and the dominion to which each god was entitled. However, because of the anthropomorphism, which dates back to before the Greeks, and perceived divine immanence, what is meant by "God" requires a much broader definition than the literal, metaphysical being that 'Creator,' 'Father,' or even Prime Mover permits. Later monotheistic figures, such as the Abrahamic formulation of God, may be held as Perfect, Omnibenevolent, Omnipotent, and Omniscient, but the enduring problem of evil, religious belief in the Devil, and (at least for some) free will that allows us to grow into the light of God does not necessarily or concisely correlate to this simultaneously held and literally believed definition. The evolution of dominating divine formulation, most notably from gods to God, and literally held ideas to more figurative expansion, has also given rise to new problems posited by definitions. This is evident in the case of the Abrahamic God compared to the Greek gods, where problems of evil, competing wills, and expansiveness were not—by necessitation—problematic or contradictory.

In addition, Greek literature also concerned the polis—or city—and how the gods were in charge of ordering the great chain of society. One need only read Virgil's *Aeneid* to see how far back the presentation of god(s) has been created in man's undivine image. 'God' is then also a constantly redefined figure used as a myth to sustain order, avoid homogeneity, frame diversity within the human portion of the great Chain of being as 'necessary hierarchy', and make sense of the world. This veers away from the less expansive, transcendental, and literal definition of God with which we started. The definition of 'God' must then be expanded to involve the human hand, which has tirelessly crafted it and cannot be limited to what was never simply a matter of the literal. It is also at this point that the figurative and literal merge. Therefore, if God is understood as a holistically figurative 'God' that has been crafted and sustained by many centuries of thinkers

and believers, what could be gained, or hoped to be gained, from His Death? In other words, how can the destruction and death of that which was designed to provide organization, explanation and, for some, even power provide any benefits, security, or desirable freedom?

To answer these questions and make sense of how the death of God may have become the equivocally accepted model of concrete 'reality', one must also understand the consequences of belief, and intellectual stirrings that took place preceding the death of God. It is one thing to manipulate the image of God to suit one's purpose, e.g., to justify hierarchy, class, or empire, but another to live religiously in accordance with His created image. Indeed, belief and believing is a conceptual division that persists today; in his *Culture and the Death of God*, Terry Eagleton notes that "in a British survey of 2011, 61 percent of respondents claimed to have a religion, but only 29 percent of them claimed to be religious" (Eagleton 1). People have been questioning what being a part of religion, or simply believing in a God, entails since the sophists of Ancient Greece, who may have inspired Sophocles' *Oedipus Rex*. One example of this figurative questioning of God and His power, control, and ability to answer life's questions was also highlighted by Dreyfus and Kelly:

[Hamlet's] famous soliloquy from Act III Scene 1, 'To be or not to be, that is the question' [(Shakespeare) quoted in (Kelly 18)], takes on the fundamental issue of whether he should choose to live or choose to die... The very idea that he understands this as a choice open to him indicates that his culture no longer takes for granted that God determines these fundamental facts of our existence... The question is whether it's a better decision - nobler in the mind - to suffer or to commit suicide. God, or the understanding of God as the divine planner of the universe, offers no help to Hamlet in considering this question... [opening] up the possibility for genuine existential questioning. (Kelly 18)

To understand, therefore, how the death of God came about, it must be recognized that 'God' is figurative not only as His image was altered by humans but also as His very existence became questionable. All the edifices and Chains that He had once been charged with building and constructing, had, therefore, borne His name as a proxy for a void in human understanding or as a contextually reasoned metaphysical estimation. The latter of these reasons is a more retrospective recognition of the figurative, a transition from being unquestioning and inconspicuously assumptive to constructing new mythology and reasoning. 'Proxy' may still be too strong a term here, given how the history of ideas is full of competing and incompatible narratives to explain or substitute explanation for 'reality' and many ideas may be equally compelling. In this sense, perhaps 'God' is not limited to being simply a proxy, but a successful myth which has enjoyed figurative rebirths, of which the religious or Abrahamic God makes up only one portion. The implications of this partiality shall be explored in greater detail in the following section. However, the current phase of philosophical thought, the figurativeness of 'God,' may itself be temporary and partial, though a return to

unequivocalness and 'blind faith' may not be the next, or any, future stage—though this is another consideration.

The intellectual stirrings and cultural movements of Romanticism, the Renaissance, and the Enlightenment that saw the virtual deification of Man replace God's ontological and moral necessity are also important in this consideration. This deification also came with the acknowledgement that 'God' is not only a non-necessity, but a non-being; the recognition that God—in the religiously understood sense—never existed. Yet, this further gave rise to the fear that with the death of God something was still being negatively lost and negatively gained. The death of God is not, in this way, a singularly positive, aspirational occurrence, but a complicated and ominous reckoning. The literal and negative loss may be the comfort one may have had with determinism, fate, and lack of control; the lesser burden on oneself if the consequences and motivations behind actions have a transcendental, a-personal explanation that evades self-responsibility. The negative gain that may be accumulated with the death of God is this burdensome responsibility and (to an extent) metaphysical and moral self-determining duty. The death of God may not have been desired even if its potential consequences, such as freedom, are desirable, as the burden of choice, responsibility, and constraints weighs heavily. Such intellectual and cultural movements may well have removed man from the great chain of being, but in its place arose new Chains which the looming, figurative figure of 'God' still embodies. The death of God, then, becomes partial, linguistic, figurative, and more reflective of our desires, intellectual evolution, and epistemological and metaphysical unease than reflective of what constitutes the literal and succinct. It is no longer a case of God or 'God,' as the title of this section states and opinion may assume, but a historical matter of God then 'God.' We will explore this further in the next section.

4. Death of God

The idea of the death of 'God' dates back to the 19th century from the works of one of the most significant figures in philosophy: Friedrich Nietzsche. A helpful description of what he meant by this idea may be summarized as: "In Nietzsche's thought, 'God is dead' does not mean 'God is now physically dead.' Rather, it is his way of saying that the idea of God is no longer capable of acting as a source of any moral code or sense of directed historical purpose" ("Death of God - New World Encyclopedia"). Indeed, the death of 'God,' which will be discussed here, is similarly not a discussion of a bodily, or literal death, but focused on the intellectual evolution, uncertainty, goals, and language that underpin it.

It is helpful to be reminded of what it is that is meant by 'God' before considering what can be hoped to be gained from His death. The previous section explored how what is meant by 'God' encompasses a constantly redefined figure used as a myth to sustain order, avoid homogeneity, frame diversity within the human portion of the great Chain of being as 'necessary hierarchy', and make sense of the world. 'God' is then a figure who stands for control, whose role may be (incompletely) understood to drain freedom away from man. 'God' as representative of "directed historical purpose" is the starting point for this essay's consideration of the concept of the death of God's appeal. Thus, godlessness may be

held as the gateway to freedom and self-definition, which provides both figurative and literal appeal. Additionally, godlessness represents apprehensive acceptance of a world in which a religious, Abrahamic God has no intellectual room or ontological weightiness.

The process by which freedom is acquired, i.e., God's death, seems to suggest that man gained the power to recognize the distinction between the literal and figurative, and that 'God'—being a figurative paradigm—may be linguistically discarded to allow for such freedom. However, the death of God—itself figurative—signals not the complete ending of a term and its unfree domination, but rather the death of a religious, transcendental, creator God, who in modernity takes on a much more figurative and expanded meaning. It also signifies the recognition of our anxiety fueled by uncertainty and vagueness, as will be explored in the next section on freedom.

The way in which these new gods have emerged, who share many of the same attributes—control, order—as those of previous ages, is demonstrated by Arthur Miller's *The Crucible*. Set in 17th century Salem during the witch trials, the play is an allegory for the United States in the 1950s, during the Red Scare (Miller and Till). The Cold War, during which the Red Scare occurred, if understood using over-simplified terms was, on the one hand, a struggle to defend American values—life, liberty, and the pursuit of happiness, individualism, difference. On the other hand, it was an attempt to quash Communism, which was, for many, anathema to these American values. In particular, Communism was viewed as an ideology marked by homogeneity, sameness, and overzealous, overreaching government as well as Big Brother and his other state-run, state-owned, state-controlled siblings. Miller choosing a theocracy in which fundamentalism, conspiracy, myth, and stifling religion took center stage as the setting upon which one may draw figurative parallels to the House Un-American Activities Committee (HUAC) suggests that post-God freedom has not yet been achieved or may even be unachievable.

In this sense, 'God' as a figurative substitute for control, myth, and power-sharing not only aligns with the ways of the religious, e.g., in Salem, but also with periods after the death of God, in which such a figurative death appears to have given rise to new systems of control, order, and power. Such a play seems to suggest that the historicized, figurative definition of 'God' still limits freedom to a figurative ideal. This also suggests that even after recognizing the figurativeness of earlier concepts of 'God,' we are in a sense unable to depart fully from all the concept encompasses. This demonstrates the mere partiality of the death of God. His replacement—the new 'gods'—may be theorized by man to substitute His absence, a further suggestion towards the anxiety and undesirability of God's death.

However, such recognition of the perhaps literal freedom that is still absent following the figurative death deals only with the blurring of meaning—literal and figurative. It does not fully account for how language is as much used to represent and estimate reality as it is to aspire towards a different reality. This different reality is what one may hope to become literal by first stating it in the figurative. In the same way, Anselm's ontological argument for the existence of God may best be understood as aspirational prayer: a figurative expression of how logic and definition ought to operate (Anselm). Expressing the death of God may

be understood as a secular prayer of how one believes the literal ought to figure in His absence. Anselm starts with the figurative in the hope of literal actualization, just as the death of God may be regarded as a figurative starting point.

If 'God' stood as a figurative expression of control and power, His death marks for some a break in the divinely ordained Chain, an escape from Oedipus' fate, the Salem "witches'" burnings, and the ability to strive towards answering genuine existential questioning (Kelly 18). The question of 'to be or not to be' in this sense, as implied by Dreyfus and Kelly, is not a matter of being with or under God, but a matter of choice, freedom, godlessness, and self-definition. Moreover, if meaning is a word's use, then the death of God was never meant to fully signal the literal (apart from literal godlessness), but rather the aspirational, and therefore, the figurative. It may be that the concept of 'God' remains, but so too does the figurative sentiment and aspiration that His' death also allows and with this, literal apprehensiveness.

What, then, is to be made of the literal freedom one may still hope to gain from the figurative death of God? Or, what is left of literal freedom, however much figurative aspiration and apprehension remain? To answer these questions, it is also worth considering how the two understandings of the death of God, i.e., its literal and figurative meaning, may be understood together. It may be true that new 'gods' and myths emerged in the old one's place and that His death is understood as figurative. In between these two distinctions lies a blurring of both figurative aspiration and the literal: a blurring of what many believe can transpire into the literal from the figurative. The history of ideas—in literature, political thought, and philosophy—has seen a plethora of attempts to offer the figurative and literal as merely temporally divided synonyms—i.e., the figurative is becoming and the literal is being. The next section traces this blurring of meaning.

Before exploring what the figurative sentiment and aspiration that accompanied His death may be, it is worth noting another reason for the incompleteness and figurativeness of the death of God. As discussed, Miller's play demonstrates that we still live with gods whose attributes—control, power, order—seem similar to the role religious gods played in pre-death societies and theocracies. Existing systems of control and power may also demonstrate the uncertainty and fear one may experience considering His death. While freedom may still be regarded as the most desirable ideal and pinnacle state of existence (as will be explored in the next section), there is an anxiety which accompanies godlessness. This fear is closely linked to the appeal secular systems of control such as strong government, security, and rights-culture may have. They are in some sense the secular gods which replaced the religious one following His death. And whilst, in turn, they too may be rejected and man may seek freedom from them, it is worth noting that there seems to be a cycle of freedom and control—both the figurative desire for and the literal presence of— that historically emerges. This cycle of systems, control, and order seems to suggest God's death was not a matter of choice but acceptance, understanding that the burden of fate and freedom lies not with God's will but with human responsibility, and ontological loneliness. Freedom then begins as multifaceted, incomplete, and historically uncertain.

5. Freedom

Aimed at being a literal actualization, freedom has been analyzed and offered through the use of the figurative. Before venturing into how this has played out in literature, political thought, and philosophy, freedom shall be defined and a distinction made. This essay discusses human freedom for two main reasons. The first is that such a freedom is more holistic, process-based, and dynamic than other types of freedom, as well as being multifaceted, leaning on other types of freedom for its actualization, i.e., self-realization. Secondly, there are certain connotations to the idea of human freedom. On a societal scale, the idea that one has the right to self-realize may be followed by the idea that one has the right to determine not only what self-realization entails, but also who gets to self-realize and how we can evaluate such distributions and actualization of rights. Human freedom then allows us to be both the literal lawgivers and judges of rights and freedom, as well as the figurative 'gods' that determine its very definition. While it does not evade the contradictions and blurring between positive and negative freedom, human freedom's holism also curbs discussion of only these tensions. There are two further distinctions to be made with regard to freedom's meaning. Maxiscule freedom, written as "Freedom," will be used to refer to what shall be shown as a figurative, aspirational ideal. Minuscule freedom, written as "freedom," will be used to refer to what may be regarded as a values reflection, example of actualization in the form of rights, or diluted Freedom. The grammar used to try and distinguish between the two is not itself indicative of the value that necessarily ought to be placed on either the figurative or literal. Such attempts to offer Freedom as an early theorization of freedom will also be analyzed, as such a consideration offers an insight into our desires and goals.

Firstly, it is necessary to further define what this Freedom means. It is assumed to mean something more specific than simply human freedom. Following a similar pragmatic genealogy for the term's meaning as attempted in the case of 'God' and His death, Hegel's principle of subjective freedom—understood in more modern terms—seems appropriate. Jon Stewart, in his essay "Hegel and the History of Human Nature," defines Hegel's use of the term as: "the idea that individuals have the right to grant or withhold their assent with regard to issues concerning right and wrong" (Stewart), adding later in the article that modernity has compounded his principle with "the elimination of rigid classes and socially determined constraints" (Stewart), perhaps indicating God's death. Broadly, this idea of self-definition, social mobility, and moral self-determination will be used to characterize Freedom as part of the previously defined human freedom. A visual model of this principle is also found during the High Renaissance when Da Vinci's *Vitruvian Man* geometrically hailed the coming together of the heavens and Earth according to man's measurements. One may say this was the preemptive, figurative completion of Michelangelo's *Creation of Adam*. Thus, the most accessible cultural references to the historical theorization of Freedom are themselves figurative, already caught up in artistic, religious, and cultural imagery. Already one can see how meaning regarding the term has long been represented as figurative, though many have held their depictions to reflect the literal. However, it is in a sense easier to discern between a figurative work of art and what is literal

than it is to discern between political and philosophical literature which often-times presents the figurative aspiration of Freedom as synonymous with literal possibility (as forced or offered by the death of God).

A paradigmatic and still bestselling example of this blurring of meaning and possibility with regard to Freedom is Ayn Rand's 1957 novel, *Atlas Shrugged*. One review from the *National Review* also highlights this blurring of meaning when it argues that "the mischief here is that the author, dodging into fiction, nevertheless counts on your reading it as political reality" (Chambers et al.). Indeed, in the same way pre-death of God literature, such as Sophocles' work, prizes its literal significance on its escape to the blurred figurative, Rand's novel bets on many of the same literary tools. The main meaning deferral that makes the book as compelling as it is figurative warrants analyzing in greater, reviewed detail:

At that point, in any materialism, the main possibilities open up to Man. 1) His tragic fate becomes, without God, more tragic and much lonelier. In general, the tragedy deepens according to the degree of pessimism or stoicism with which he conducts his hopeless encounter between human questioning and the silent universe"... Or, 2) Man's fate ceases to be tragic at all. Tragedy is bypassed by the pursuit of happiness. Tragedy is henceforth pointless... Henceforth man's fate, without God, is up to him, and to him alone. (Chambers et al.)

The godless and free man, in Rand's characterization, seems at least in some sense tangible. It is, however, her use of a black-and-white morality, unrealistic storylines, and appropriative use of mythology which makes it so puzzling as to how one may draw literal parallels to this figurative work. Our apprehensive reluctance to "shrug" (Rand, chap. III) as her protagonist does, escaping the new 'gods' also does not make her Freedom manifesto more relatable or appealing. It is a *Guardian* article fruitfully titled "Atlas Shrugged is Absurd but Strangely Compelling" that cements this:

Rand's world is a place of black and white morality, good and bad people and absolutely no shades of grey. Consequently, none of the characters or storylines are at all believable... turning to Atlas Shrugged because you don't like the way things are going is the equivalent of diving for the centre of the fire because the frying pan got too hot. (Jordison)

Despite this world of "black and white morality, good and bad people and absolutely no shades of grey," there are many readers who are still enamored by this book. In a *New York Times* review, one interviewee who named their company after John Galt recalls "new clients usually ask: 'Hey, where is John Galt? How come I'm not important enough to rate a visit from John Galt?'" (Rubin). Whilst the clients may not be familiar with the novel, the literality with which they took the figure of John Galt is no less serious than some who believe that the Freedom espoused by Rand can literally exist. This Freedom is embodied by one

of her characters, John Galt, and readers who interpret her text may literally see themselves alongside Galt as Freedom's existents. In this sense, literal and ideological exponents of her work—and other works of the same figurative kind—may find eisegetical appeal in the way in which such storylines offer aspirations and express desires. Interpretation and meaning are not concrete, reductionist or Platonic and so it is even easier for us to find literal aspiration in figurative language. Our blurry interpretations may then persist without warranting complete interpretative refutation. This blurring and purposeful merging of what one believes to be the figurative and literal into the singularly literal is exemplified by Rand's novel, especially through the use of the figurative, to give Freedom and power to man.

The title of her novel, *Atlas Shrugged*, is perhaps the least critiqued or referenced part of the book, both by those who literally believe in its message and those who think it is figurative. The title is a reference to the myth of Atlas, who held heaven and earth apart" (Atlas | Greek Mythology"). However, Rand's use of the myth is to demonstrate that instead of bearing the weight of the world, the "producer" (Rand), or the rational, capable man, must shrug it off and neglect social responsibility, favoring atomism and rational egoism. This is shown in a scene between two of her characters:

'If you saw Atlas, the giant who holds the world on his shoulders, if you saw that he stood, blood running down his chest, his knees buckling, his arms trembling but still trying to hold the world aloft with the last of his strength, and the greater his effort the heavier the world bore down upon his shoulders—what would you tell him to do?'

'I...don't know. What...could he do? What would you tell him?'
'To shrug'. (Rand, chap. III)

Rand may not have been implying that the literal earth is on the "producer's" shoulders. However, there is an element of literality with which she employs this myth: The values which Rand holds—materialism, wealth, power, Freedom, selfishness—are, to her, literal. The social responsibility and burdens which Atlas bears on his shoulders are also literal. Thus, Rand recognizes in this metaphor the literal and intends for her audience to see the same, and for those worthy among them, like Atlas, "to shrug". Most interestingly, the burden of Atlas, and that of the capable man, seem similar to the burden taken on by Jesus in the Gospel of Matthew before man was able to gain salvation and a chance to enter heaven, perhaps a religious shrug of past (original) sin. This demonstrates the parallel between the way in which the figurative, including myth and metaphor, has been used to characterize 'gods' both in traditional religious discussions and post-God, Freedom discourse. Both cases also sustain a significant following and literal interpretation. More deeply, the parallel between the shrugging of Jesus on behalf of mankind and the shrugging of Atlas seems to be an expression of desired freedom, from sin and responsibility respectively, and yet both seem incomplete. Sin seems to have continued and salvation is still decidedly (for some theologians and interpreters) uncertain, and responsibility, 'gods' and 'chains' of control,

seems to, despite Rand's attempts, continue binding us. Both seem, in this sense, confined to some figurativeness and questionability.

The almost untraceable leaps between what is held to be the figurative and literal mean that the author and the audience may both believe that Freedom and other figurative idealizations are literal, and attainable, despite unchanging social constraints. The *National Review's* review cements this: "Lest you should be in any doubt after 1,168 pages, she assures you with a final stamp of the foot in a postscript: 'And I mean it'" (Rand qtd in Chambers et al.,) (Chambers et al.). There can be little doubt about the certainty with which Rand and many others hold this claim. Such a novel demonstrates how easy it is to blur the boundaries between the literal and figurative. For some it is the literal where such theorizations, ideas, and novels appear to be a blur of the figurative, and for others it is the literal where the figurative is only that which preempts, expands, and aspires towards that which may well constitute, or already constitutes, the literal.

If there exists a division between those who determine meaning, what can we collectively call freedom? What rights or conventions may be hailed as the actualization, and dilution, of Freedom or human freedom? Though much less expansive than many formulations offered by thinkers such as Rand's social atomism, one cannot necessarily claim that modernity is without any sense of heightened liberation or freedom. This may be especially true if freedom is defined as antithetical to 'God,' i.e., some form of self-definition, social ability, and lack of given social order. However, as some everyday forms of freedom that have coincided with the rise of modernity seem standard, trivial, and unideal, it can be easier to focus on our desires for more Freedom (in the form of freedoms). In his same essay on Hegel, Stewart shares a similar idea that whilst "the emergence of subjectivity in the ancient and medieval world was a liberating development, in the modern world the pendulum has swung too far to the opposite extreme" (Stewart). One can interpret this to mean that we still desire figurative, idealized Freedom and wish for our literal freedoms, in the forms of rights and conventions, to reflect this ideal. And so, the remainder of this section shall, therefore, focus on how freedom finds its origins in figurative Freedom and is constantly critiqued from its viewpoint.

There have been rights established in the form of declarations, constitutions, and laws. Rights and freedoms such as choice and individuality are celebrated as markers pointed towards a Freer, unchained, and self-defined man. However, such examples of rights find their origins, as much of political thought does, in abstract philosophical thought that veers markedly clear of everyday freedoms. Preferring ideals or Freedoms that cannot be pragmatically carried out, at least not in any current society, philosophical abstractions regarding rights such as uninhibited freedom of speech or a complete lack of social responsibility remain a part of figurative ideals. It is in a sense due to their abstract origination that freedom may be viewed as conceptually lesser than Freedom and not fully aligned with possibilities opened up by the death of God. In this way, God's death and Freedom both inform literal freedoms and are seen as worthy, and inform partially literal standards that the literal is to be held against. The intrinsic value of such standards is not necessarily negative. However, deferring meaning from the literal to the figurative and vice versa in order to express our desires may sug-

gest that language causes, and is caused by, aspirational and epistemological uncertainty. We are unsure of how far our Freedom and freedoms extend and whether we actually desire such an uninhibited state of being.

The ways in which rights and freedom represent incomplete Freedom reflect two ideas: 1) the fear of God's death and 2) the appeal of constrained freedom. The epistemological (and metaphysical) uncertainty that underpinned much of belief in God means that in practice, upon His death, there is less appeal in man signaling the coming together of the heavens and earth. The *Vitruvian Man* freed from the creation Chain becomes less of an emphatic desire and more of an apprehensive acceptance. The Hobbesian picture of the state of nature seems to be lacking not a 'Leviathan' (Hobbes and Brooke 54) but a 'God,' religious or figurative, that prevents the freedom of man from devolving into self-destruction. This picture is, in a sense, what one may visualize as coming with the death of God and unconstrained freedom. This is also where constraints come in. Rights and freedom which weakly reflect and actualize Freedom also involve security and negative freedom, e.g., the right not to be discriminated against and the right not to have one's beliefs questioned. There are also collective rights which may be understood as protection against the chaos and Hobbesian state of nature that absolute individual rights and unconstrained freedom may lead to. In this way, if the same figurative definition of 'God' is used, not only does this exemplify the partiality of the death of God but also the way in which figurative Freedom may distort what one may literally (in practice) desire and/or regard as freedom; it also shows that we remain apprehensively under figurative 'gods.' By viewing this thought transition from God to 'God' we are able to trace and analyze the nihilistic concern one may have in the absence of the divine. Freedom in this exploration is an uncomfortable bedfellow, quashed and confined, at once idealized and unwanted, critiqued, and compared to systems of control with which (in terms of security and certainty) it may not compare.

6. Setting the Discussion and Review

Thus far, how one defines 'God' has been explored alongside His death, as well as its potential consequences, i.e., godlessness and Freedom and accompanying apprehensions. Such consideration has exposed how meaning, i.e., language's use, as well as how boundaries between the literal and figurative, are changed, blurred, and deferred in an effort to make greater sense of the world, either as an estimator of concrete, certain knowledge, or as a figurative aspiration towards that which the literal ought to strive. Whilst necessary to construct the setting in which the discussion may continue, the underlying, somewhat implicit or inconspicuous motivations and sentiments that fuel such blurring have not been fully linked to its epistemological setting. The deeper unease and uncertainty which grounds thought briefly highlighted in the first and preceding section, i.e., the potential reason beyond our blurring of meaning, our desires and fears, will be the focus of this section.

As discussed, the definitions of 'God' and Freedom are debatable. This is true not only of 'God's' definition but of most definitions because they are open to

change, evolution, adaptation, and confusion, as our desires, knowledge, and linguistic tools also change. Debatable definitions are why adopting a widely Wittgensteinian approach to language, understanding it through its use, may be the most pragmatic methodology (“Ludwig Wittgenstein (Stanford Encyclopaedia of Philosophy)”). This approach is also an argument against a removed, lofty realm of unchanging, perfect Forms where definitions are concretely eternal. Unchanging definitions are an originally Platonic conception which remain a linguistic misconception of what definitions and meanings may be. Definitional evolution adds fuel to the blurring of the literal and figurative because, as seen in the case of Hamlet and God, what was once held to be the literal, with a concrete, eternal definition—like that of the Abrahamic God—is subject to a transition to the figurative upon one’s changing interpretation and/or understanding of His existence and/or existential necessity.

Therefore, goals, desires, and uncertainties—originally aspirational and blurred—also allow language to more easily slide between the literal and figurative. When such desires become more aspirational, abstract, or unrealistic, meaning is more strongly blurred in an attempt to accommodate our new sense of godlessness, Freedom, and self-definition. This suits the purposes of goal-making, but diminishes what we may be capable of as seen in the case of God’s Death, Freedom, and what fears and apprehensions may arise from these considerations. Nevertheless, understanding language in terms of its use also allows for a better understanding of what we aim for and why we constantly turn to the figurative. This is not to say that literal goals, which may be more pragmatic, are any more suitable, appealing, or ethically desirable than that which occupies the abstract; this is not an evaluation of moral value. Nevertheless, it is still possible to see how it may be more comforting, due to both the perceived limitations of the literal and the endlessness of the figurative, not only to make meaning figurative, but also to aim for the figurative in a very literal sense. Thus, language provides a vehicle by which desires and aims take flight—if only one of figurativeness and fancy—as exemplified in the consideration of godlessness and freedom, escaping the perceived banality or uncertainty of the literal. In the first section on the literal versus figurative, discerning between the two was shown to be difficult, especially because such a task can involve value judgements, which have now been exemplified by the incomplete life cycle of God and ‘God.’

Such figurative appeals are, as noted, both symptomatic and, in part, consequently prescriptive of the vague literal. This is not to state, however, that the figurative is any clearer. The deeper unease and uncertainty we face breeds, and is bred by, the blurring and deferral of meaning which dominates the history of ideas—a linguistic feedback loop. The literary examples explored—Sophocles, Shakespeare, Miller, and Rand—all involve the figurative and link it to the literal with an apparent moral, aspirational, and impassioned gaze reflecting the battle against uncertainty. This may be attributed to the way in which epistemology is limited and imprecise, and how our understanding of life is vague, fragmentary, incomplete, and contingent. This is reflected in intellectual evolution as well as language, such as the death and redefinition of ‘God’ explored earlier. These changing definitions, goals, and epistemologies, as well as constant debates about the existence of God, are not themselves indicative of the certainty of God’s exist-

ence. However, their long-standing presence is a critical revelation about our relationship with language and the world around us. This paper aimed to demonstrate the extent to which meaning slides between the different, and for some not so different, dimensions of meaning and our understanding. In particular, the topics of God, His death, and Freedom the struggle between meaning and knowledge. Moreover, whilst there have been many thinkers who argue fervently that God must, cannot, may, or even ought to exist, they have all done so using language—both literal and figurative—as an expression of their espoused beliefs. This paper has also aimed to posit the underlying tensions and causes of our relation with language.

Although an attempt was made to distinguish between the literal and figurative in the introduction to this essay, through the considerations of the ideas discussed, this distinction warrants revisiting and reviewing. Context and intention were offered as two ways of discerning between literal and figurative meaning, with regard to how language may be said to reasonably correspond to perceivable reality, what the author intends, and how one may interpret language's meaning. Despite this, and other meticulous attempts to differentiate between them, there is a reasonable case to be made for the at least partial figurativeness of all language and meaning. As seen with the redefinition and figurativeness of 'God,' 'His' definition relies on reference to other definitions. What makes possible the figurative and linguistic death of God is an understanding of what the absence of God means and entails. Thus, this is also a claim against a reductionist view of language, as the history of ideas, though focalized in this paper, shows the necessity of holism and the way in which there is a conspicuous and inescapable diversity to both how one may define a term and how one may interpret and/or understand definitions. Thus, the considerations in this paper have also provided a way in which one may better understand the development and evolution of language. This way of analyzing language may be most insightful, even against a background of uncertainty and vagueness which may make static, concrete definitions all the more appealing. This paper's discussion of the literal and figurative also demonstrates how fragile and blurred the boundaries, and therefore, definitions, are between the two. In so doing, language and perceived reality become entangled in a loop of the literal and figurative whereby the two dimensions of meaning are exchanged as thought evolves and narrative explanations compete.

7. Objections and Responses

There are, of course, objections to the claims made, and potential disagreement over the way in which God' and Freedom have been defined. There are also arguments against the epistemological and linguistic claims made. This final section shall, therefore, explore and respond to these objections not only to dialectically consider alternatives to the claims made, but to hopefully also strengthen the arguments put forward.

Earlier, 'God' was used to refer to a constantly redefined figure used as a myth to sustain order, avoid homogeneity, frame diversity within the human portion of the great Chain of being as 'necessary hierarchy', and make sense of the world. Figurative 'God' is a considerably expansive meaning which might still

evade what religious believers actually intend to worship. It may then seem presumptive, given our 'post-God' age, and appropriative, given the lack of religion, for 'God,' including the formerly defined God, to be given such a nonreligious definition. This is the first objection that ought to be considered. It may be argued that despite the historical questioning of God's existence, there is a certain level of hindsight and modern-secular bias in associating godlessness with Freedom or even morality; and 'God' only with systems of power, control, and myth. However, the history of ideas, when viewed retrospectively, allows us to make sense of the somewhat common sentiments towards God and what belief entailed. It also allows us to analyze what God, or 'God,' entailed on a societal level—from theocracy to divine rights to the justification of rigid social structure. It is, of course, noted that 'God' also lacks the moral sentiments and epithets that one may wish to attach to a religious God: Omnibenevolent, Father, Creator, Saviour. However, the literature from which such epithets are derived may neglect to cement this claim in any literal or definite sense. For example, in the Old Testament, God's attributes lean more towards 'Grand plan' benevolence and cruel-to-be-kind decrees than towards a valid objection to such contrary epithets' omission. The response to this objection also highlights that concepts, even when theorized in such a way, cannot be justly analyzed or considered ahistorically or atemporally, for, as previously explored, language and thought is a process that evolves, given changes in ideas, cultures, and language itself. Thus, the supposed lack of religiosity in the figurative 'God's' given definition accounts for this holism and process.

There may, however, be a sub-critique raised regarding this essay's definition of 'God.' It may be argued that such a definition and meaning attachment moves the goal spot and "begs the question", making 'God' somewhat synonymous with power, order, and control, inevitably leading to the figurative and literal becoming blurred. This would, therefore, suggest we lack certainty or clarity, as constant redefinition and changes in power structure occur. Although such a definition leads to this essay's conclusion, if one looks at the meaning of a word as its use, then it is not "begging the question", in terms of the traditional philosophical fallacy, but rather offering a descriptive analysis of what is entailed when we say 'God' or even refer to a religious God in the most fundamentalist, hermeneutically literal sense.

Another objection that may be anticipated is against the claim that knowledge and epistemology are uncertain or vague. Although it may stem from the same objection against a non-religious, changing definition of God, it is worth considering as a separate objection. One of the central premises upon which the conclusion of this essay is based is the setting of vagueness and uncertainty, which increases the blurring of meaning. It may be argued that the end of inquiry is not only achievable but achieved in the history of ideas, which has given rise to a competent understanding of reality and language that allows us to avoid figurativeness and, in any event, distinguish between the literal and figurative. However, etymology provides a strong challenge against this objection. For example, a few years ago, the ironically extended definition of the word literally divided academic and popular opinion. Although there were a few tangents to the debate, the central question was, "Is it ever okay to use literally to mean figuratively?" (Edi-

tors of Merriam-Webster). Some felt the word had been given a redefinition inverting the literal meaning of *literally* to be purely figurative, an assumed linguistic nightmare. It was not long until several articles in the *National Geographic* (Skurie), *Thought Co* ("What Literal Meaning Really Means"), and even some dictionaries detailed the long use of the word in the figurative sense, though its frequency of use had certainly proliferated. Similarly, the word *idiot*, has departed from its originally political meaning to be applicable, or at least applied, in numerous circumstances. This is not a rare occurrence in language or etymology, especially as old words stop being used and new and partly inherited words come to dominate. This constant evolution reveals that our use of language is not definitive and changes as we ourselves change (and vice versa). As previously argued, there is also no higher realm to which we can climb (Plato's Symposium) in order to understand, appreciate, or grasp definitions in their eternal and unchanging Forms. Thus, the premise of uncertainty finds its genealogical hallmark in language, providing a reasonable basis upon which the arguments have been concluded. Similarly, the previously discussed holistic perception of language also provides a compelling case against this objection.

Finally, in the introduction, the intellectual positioning of this research was mentioned as worthy of critical review. From current and certain perception, this paper is contemporary and can attempt retrospective analysis with the full, or in this case selective, weight of intellectual and historical endeavors. The current post-god age also provides a subjectively curated lens through which to view cultures where unquestionable belief dominated, i.e., in some pre-modern and ancient societies. The literature explored in this paper has also been purposefully chosen in order to best demonstrate the claims made. Both of these factors, i.e., intellectual positioning and literature used, are arguably as indicative of the conclusions reached as the arguments made. These factors are recognized and this paper's principal claim is accordingly not one of Truth, Enlightenment, code-cracking, or necessary singularity, but of one potential and plausible narrative that allows us to trace the history of ideas. The premises are not definitive, but in this research's context, they are the ones with which one may find the most validity, however temporary or partial. The point of reviewing this paper's structure and content is to provide meta-textual cementation of the paper's claims: thought and language evolve as competing narratives clash and intellectual mediums and disciplines attempt to make sense of the world. After all, whilst the transition from God to 'God', His death, and then freedom could not be made perfectly concise or narrow by this paper, the literal intellectual changes and loops between these transitions are themselves fraught with tension, disagreement, uncertainty, and battling narratives.

8. Conclusion

Sophocles' *Oedipus Rex*, Shakespeare's *Hamlet*, and Rand's *Atlas Shrugged*—and their eponymous characters (and John Galt)—are few of the many characterizations that capture a historical narrative. The authors' intentions and corresponding perceived reality need not have intersected, or even been known, for these literary figures to have been pivotal markers of an intellectual and cultural evolution. This paper has, therefore, sought to construct and present an argument for a potential explanation behind the blurring of meaning—literal and figurative—and the intellectual evolutions—God, His death, and figurative 'rebirths'—using literature as a demonstrative medium. As previously stated, the genealogy in this paper is one of many potential narratives that may be attributed to the topics of godlessness and freedom. Nevertheless, the epistemological, linguistic, and cultural backgrounds offered—uncertainty, apprehensive acceptance of God's Death, linguistic change and the incomplete transition to Freedom and freedom—strengthen the arguments presented. The main point of this paper, therefore, was to provide a plausible explanation and genealogy for the conceptual division to which language and the history of ideas have given rise, and analyze our place and involvement within these phenomena.

Indeed, our endeavours are fraught with the blurring of the literal and figurative which allows us to manipulate, unknowingly and purposefully, the boundaries of meaning on the account of context, intention, and interpretation. This treatment of language is not new or unique to English and has been given a substantial and well-represented place in the history of many ideas, across a variety of intellectual disciplines and genres, extending even into the visual arts. Though focalized and niche through the evaluation and analysis of language with regard to 'God,' God's Death, and Freedom, this discussion is not itself new to the field of philosophy. Nevertheless, continuing and revitalizing the discussion allows us to not only critically assess the way in which we use language and consider meaning, but also reflect upon the vagueness and uncertainty that sets the background for our desires, aspirations, language, and thought. There is, of course, much more to be explored and many questions not yet asked or answered: Can the two dimensions of meaning be pried apart? Does this matter? Could we return to a state of unquestioned belief in God, or polytheism? Or, does our very questioning of the discussed topics reveal or suggest a tendency, blurring, or inclination of its own? Thus, this essay may also serve as a starting point from which new ideas, interpretations, narratives, claims, and considerations may stem and evolve.

Works Cited

- Anselm. *Proslogion: Discourse on the Existence of God*. Beloved Publishing LLC, 2014.
- Atlas | Greek Mythology." *Encyclopedia Britannica*, 5 Aug. 2021, www.britannica.com/topic/Atlas-Greek-mythology.
- Ayn Rand | Biography, Books, & Facts." *Encyclopedia Britannica*, 2016, www.britannica.com/biography/Ayn-Rand#ref142609.
- Case, Holly. Left and Right Are Both Gripped by an Identical Fear: Homogeneity | Aeon Essays." *Aeon*, 28 Dec. 2018, aeon.co/essays/left-and-right-are-both-gripped-by-an-identical-fear-homogeneity.
- Chambers, Whittaker, et al. Big Sister Is Watching You." *National Review*, 4 Oct. 2019, www.nationalreview.com/2005/01/big-sister-watching-you-whittaker-chambers.
- Death of God - New World Encyclopedia." *New World Encyclopaedia*, 2021, www.newworldencyclopedia.org/entry/Death_of_God.
- Eagleton, Terry. *Culture and the Death of God*. Reprint, Yale University Press, 2015.
- Editors of Merriam-Webster. Did We Change the Definition of Literally?" *The Merriam-Webster.Com Dictionary*, 27 Aug. 2019, www.merriam-webster.com/words-at-play/misuse-of-literally.
- Hobbes, Thomas, and Christopher Brooke. *Leviathan (Penguin Classics)*. First Edition, Penguin Classics, 2017.
- Jordison, Sam. Atlas Shrugged Is Absurd but Strangely Compelling." *The Guardian*, 21 Aug. 2019, www.theguardian.com/books/booksblog/2009/mar/27/ayn-rand-atlas-shrugged.
- Kelly, Hubert Dreyfus, Sean Dorrance. *All Things Shining: Reading the Western Classics to Find Meaning in a Secular Age by Hubert Dreyfus Sean Dorrance Kelly*(2014–10-07). Free Press, 2021.
- Ludwig Wittgenstein (Stanford Encyclopedia of Philosophy)." *Stanford Encyclopaedia of Philosophy*, 2 May 2018, plato.stanford.edu/entries/wittgenstein.
- Miller, Arthur, and Lawrence Till. *The Crucible*. Heinemann Educational Books Ltd, 1992.
- Plato s Symposium*. University of Chicago Press, 2001.
- Nietzsche, Friedrich. *Thus Spoke Zarathustra*. Independently published, 2020.
- Protagoras (Stanford Encyclopedia of Philosophy)." *Stanford Encyclopedia of Philosophy*, 8 Sept. 2020, plato.stanford.edu/entries/protagoras.
- Rand, Ayn. *Atlas Shrugged*. 1st Edition, Penguin Books, 2007.
- Rubin, Harriet. Ayn Rand s Literature of Capitalism." *The New York Times*, 15 Sept. 2007, www.nytimes.com/2007/09/15/business/15atlas.html.
- Skurie, Jaclyn. The Literal Truth About the Word Literally." *Culture*, 3 May 2021, www.nationalgeographic.com/culture/article/1308016-words-literally-oxford-english-dictionary-linguistics-etymology.

- Sophocles, et al. The Three Theban Plays: Antigone; Oedipus the King; Oedipus at Colonus." *Oedipus the King*, 1st ed., Penguin Classics, 2000, pp. 129–251.
- Stewart, Jon. Hegel and the History of Human Nature | Aeon Essays." *Aeon*, 2 Dec. 2019, aeon.co/essays/hegel-and-the-history-of-human-nature.
- Turbayne, Colin. *The Myth of Metaphor*. Amsterdam, Netherlands, Amsterdam University Press, 1970.
- What Literal Meaning Really Means." *ThoughtCo*, 2020, www.thoughtco.com/literal-meaning-language-1691250.



Race in the Eugenics Movement and the Progressive Era, 1910-1939

Ariane Vidal

Author Background: *Ariane Vidal grew up in the United States and currently attends San Francisco University High School in San Francisco, California in the United States. Their Pioneer research concentration was in the field of Science, Technology, and Society and titled "Modern Technology: The Future, Present, and Past."*

Abstract

This study investigates the role of race in the eugenics movement and its early twentieth-century critiques. It further draws attention to the role of race in other Progressive Era movements in relation to eugenics. Eugenic organizations' publications in addition to studies written by pioneers of eugenics were analyzed to learn how race impacted the field's trait organizational systems and scientific evidence. Moreover, scholarly critiques of eugenics and writings by social activists in Progressive Era movements were explored. This led to the understanding that science and race informed each other in the eyes of not only the eugenics movement's supporters, but also the field's early twentieth-century critics. The eugenics movement collapsed post-World War II due to efforts in anti-racism; however, prior critics only found fault with the methodological aspects of eugenics. Race further seeped into the ideologies of Progressive Era activists we today associate with positive reform, such as birth control and labor activists, which exhibit ties to eugenic thought.

1. Introduction

Francis Galton, an English statistician, first coined the term "eugenics" in 1883 in his book, *Inquiries into Human Faculty*.¹ He described the belief that human traits (whether they be physical, mental, or emotional) were inheritable.² Eugenics, in turn, revolved around the desire to improve the human race by increasing the number of what were deemed high-quality traits in the human population and decreasing the number of what were deemed low-quality traits. By the turn of the century, scientists began applying Mendel's Laws of Inheritance, formerly used mainly in cases of plants

¹ Francis Galton, *Inquiries into Human Faculty* (New York: Macmillan, 1883), 17.

² Garland E. Allen, "The Misuse of Biological Hierarchies: The American Eugenics Movement, 1900-1940," *History and Philosophy of the Life Sciences* 5, no. 2 (1983): 108.

and animals, to human traits to give eugenic philosophies a biological foundation.³ The eugenics movement rose to popularity in Western Europe and the United States, reaching its peak in the 1920s.⁴ The sources of social problems were seen as rooted in genetics, inherited through families like eye color or hair color, rather than rooted in societal circumstances.

Ideas of race pervaded the eugenics movement, though the ways in which they did may seem unexpected. Race played a role in the distinctions between low and high-quality traits and individuals; however, eugenic organizations' methods of data collection often did not involve obtaining information regarding one's race. Charles Davenport, a well-known eugenicist of the early twentieth century, advocated against the production of offspring from two parents of different races, not explicitly because of a superiority of one race over another, but because of a supposed unpredictability of traits in the offspring.⁵ Furthermore, the collapse of the eugenics movement is often associated with the end of World War II, though critics of the field existed well before then. What separated them from the critics post-World War II was that they still believed in eugenics' goal of improving humanity; they only found fault within the methodology of certain eugenic organizations. In the present day, we often associate eugenics with the Nazi Party, yet the majority of the sterilization operations conducted in the United States took place during World War II.⁶ Only as Adolf Hitler was defeated did the American scientists and authority figures "embrace[e] cultural explanations of human difference."⁷

The Progressive Era lasted from 1897 to 1920 in the United States, bringing about social reform in labor, women's rights, and public health, among other areas. The eugenics movement rose to prominence during the second half of the Progressive Era, around 1910. Hence, many Progressive Era activists worked alongside eugenicists. Margaret Sanger, a notable birth control activist during this time, for example, was a eugenicist herself.⁸ Moreover, race played a significant role in determining who had access to resources such as birth control. Today, Progressive Era movements are associated with positive social change, as they aided in the improvement of human living conditions. The eugenics movement had a similar goal; however, the period is well-known for constituting a horrific time in history. Examining eugenics' role in the greater Progressive Era will allow for a reconsideration of the values of the movements our current society regards as reformative.

Race was embedded into eugenic thinking of the early twentieth century, influencing trait organizational systems, biological evidence for these systems, and critiques of the movement itself. Furthermore, the eugenics movement coincided with the Progressive Era; therefore, much of the social reform we celebrate today has ties to twentieth-century eugenic thought.

³ Garland E. Allen, "The Eugenics Record Office at Cold Spring Harbor, 1910-1940: An Essay in Institutional History," *Osiris* 2 (1986): 228.

⁴ Allen, "The Misuse," 112.

⁵ Jayne O. Ifekwunigwe, ed., *'Mixed Race' Studies: A Reader* (n.p.: Taylor & Francis Group, 2004), 102.

⁶ Alexandra Minna Stern, *Eugenic Nation: Faults and Frontiers of Better Breeding in Modern America*, second ed. (n.p.: University of California Press, 2016), 17.

⁷ Stern, *Eugenic Nation*, 3.

⁸ Dana Seitler, "Unnatural Selection: Mothers, Eugenic Feminism, and Charlotte Perkins Gilman's Regeneration Narratives," *American Quarterly* 55, no. 1 (March 2003): 67.

2. Historical Framework

One of the most prominent eugenic organizations was the Eugenics Record Office (ERO). The ERO was founded in 1910 by Charles Davenport, who established the organization's base at Cold Spring Harbor Laboratory in Long Island, New York.⁹ At the time of the ERO's founding, Davenport called upon Harry Laughlin to work as his superintendent. The office was largely funded by Mary Williamson Harriman, the widow of telegraph and radio executive Edward Henry Harriman, for the first seven years of its operation.¹⁰ Davenport also received donations from the public, as he advertised the ERO as an organization dedicated to improving humanity.

The Eugenics Record Office published their first report in June 1913, detailing the goals for the organization. They were as follows:

1. *To serve eugenical interests in the capacity of repository and clearinghouse.*
2. *To build up an analytical index of the traits of American families.*
3. *To study the forces controlling and the hereditary consequences of marriage-matings, differential fecundity, survival and migration.*
4. *To investigate the manner of inheritance of specific human traits.*
5. *To advise concerning the eugenical fitness of marriages.*
6. *To train field workers to gather data of eugenical import.*
7. *To maintain a limited field force actually engaged in gathering data for eugenics studies.*
8. *To co-operate with other institutions and with persons concerned with eugenical study.*
9. *To encourage new centers of eugenics research and education.*
10. *To publish the results of researches and to aid in the dissemination of eugenical truths.*¹¹

These goals centered around two common themes — the ERO as a data repository site and the ERO as an organization dedicated to advocacy for the production of fit offspring. These goals manifested themselves in the Eugenics Record Office's actions throughout their years of operation from 1910 to 1939. The ERO engaged in data collection through the distribution of surveys to the American public and the act of field work by the organization's employees. Surveys and interviews by the ERO typically asked individuals questions regarding the physical and mental traits exhibited by their families. This allowed ERO employees to make conclusions about what they saw as the inheritability of these traits throughout the family's generations. This data was stored in the ERO base at Cold Spring Harbor Laboratory. Additionally, surveys such as "Index to the Germ Plasm" related to the goals regarding the promotion of the production of high-quality traits. This survey, in particular, allowed for the ERO to judge the quality of traits in the theoretical offspring of two parents.¹²

The scientific justification of eugenics lay within the application of Mendel's Laws to human traits. Davenport utilized the concept of dominant and recessive genes, arguing that parents with no signs of a trait such as "mental defectiveness"

⁹ Allen, "The Eugenics," 227.

¹⁰ Allen, "The Misuse," 234.

¹¹ Harry H. Laughlin, *Report No. 1* (n.p.: Eugenics Record Office, 1913).

¹² Robert M. Yerkes and Daniel W. LaRue, *Outline of a Study of the Self* (n.p., 1913), (21).

could produce offspring exhibiting this trait.¹³ According to Mendelian theory, this “mental defect” appeared to be a recessive gene that negatively impacted mental development. Scientists at the Eugenics Record Office used Mendel’s Laws to determine how traits would supposedly be inherited in a family, further using the laws as evidence in advocating against certain pairings of parents.

Eugenicists used their data as a justification for reforming the human population through sterilization, immigration laws, birth planning, and the enforcement of moral duty. Throughout the eugenics movement, thirty-two US states passed sterilization laws, and over 60,000 people were sterilized.¹⁴ Moreover, the Johnson-Reed Immigration Act was passed by the US government in 1924 as a means of restricting the immigration of what were deemed undesirable European racial and ethnic groups.¹⁵ The act further outlined that those deemed “feeble-minded” by a scale designed by Lewis S. Terman, a member of another eugenic organization, the American Eugenics Society, would be sentenced to compulsory sterilization.¹⁶

The eugenics movement occurred as part of the Progressive Era in the early twentieth century. The Progressive Era had a significant impact on the American economy, introducing regulations on working conditions, banning child labor, and setting a minimum wage.¹⁷ Movements regarding issues in women’s suffrage and birth control also came about during this period. Throughout the Progressive Era, scholars investigated the roots of social problems, using scientific inquiry to pose solutions for them.¹⁸ Similarly, eugenicists promoted a positive outlook on their actions by emphasizing the overall goal of obtaining a society deemed fit by scientific evidence.

3. Categorization Systems

The Trait Book, included in the Eugenics Record Office’s sixth bulletin, featured a trait organizational system. *The Trait Book* was intended to be used by employees to trace inheritable mental, physical, and social characteristics throughout a family tree. Each trait was associated with a series of numbers. For example, traits regarding the brain and spinal cord were classified with the number thirty-one, and traits regarding special abilities, such as reading and spelling, were classified with the number forty-five.¹⁹ In the *Journal of Nervous and Mental Disease*, published as part of the Eugenics Record Office’s fourth bulletin, Dr. H. Valentine Wildman Jr. described a system used to categorize mental ability, in which the term “idiots” was used to describe the group with the lowest ability.²⁰ Labels such as cretinoid, a term describing foolishness; hydrocephalic, a term describing a condition in which one’s brain fills up with fluid; and microcephalic, a term describing small brain size, were

¹³ Davenport, *Heredity in Relation*, 66-67.

¹⁴ Stern, *Eugenic Nation*, 115.

¹⁵ Allen, “The Misuse,” 119.

¹⁶ Stern, *Eugenic Nation*, 19.

¹⁷ “Retrospectives: Eugenics and Economics in the Progressive Era,” *The Journal of Economic Perspectives* 19, no. 4 (Fall 2005): 213-214.

¹⁸ “Retrospectives: Eugenics,” 217.

¹⁹ Charles Davenport, *The Trait Book* (n.p.: Lord Baltimore Press, 1912), 3.

²⁰ H. Valentine Wildman, Jr., “Psychoses of the Feeble-minded,” *The Journal of Nervous and Mental Disease* 42, no. 8 (1915): 530.

used as criteria for this category of people. When field workers collected data, they would record information about the various mental and physical traits of a family on index cards and, using the number system in *The Trait Book*, would sort them into folders and boxes at the Cold Spring Harbor Laboratory.²¹ A field worker would also be able to cross-reference a trait according to the family's surname, location, and classification number. This data would be used to make conclusions about the relationships between certain characteristics of a family.

Ideas of race seeped into what was regarded as a scientific categorization. Several other criteria to be deemed an "idiot" in the *Journal of Nervous and Mental Disease* included belonging to ethnic groups such as Mongolians, or descendants of those living in the East Asian country of Mongolia; Negroids, or descendants of Black racial groups in Africa; and American Indians, or descendants of indigenous peoples of the United States.²² The use of intelligence quotient tests, most commonly the Binet-Simon test, were only implemented to determine the mental ability of persons without one of these characteristics.²³

Along with field work, the Eugenics Record Office used questionnaires as a mode of collecting data.²⁴ Dr. Robert M. Yerkes and Dr. Daniel W. LaRue's *Outline of a Study of the Self* promoted a method of self-analysis in determining one's overall fitness in society as well as the fitness of their future offspring. The outline was divided into three parts: "The Ancestral History of the Self," "The Development or Growth of the Self," and "The Self of Today."²⁵

"The Ancestral History of the Self" asked the individual completing the outline to construct their own family tree in the format of a pedigree, or a diagram of inheritable traits, tracing the presence of what they believed to be notable traits in their family and using symbols to indicate how the traits had manifested themselves throughout generations. An example from the Eugenics Record Office was provided as a template; this pedigree included the presence of traits such as "natural leader of men" and "remarkable memory" throughout a family.²⁶ This section of the outline also included a questionnaire entitled, "Record of Family Traits," which requested information regarding the individual's education, occupation, history of illness, and special abilities, among other qualities.²⁷ The individual needed to provide the same information about their parents, siblings, children, spouse, parents-in-law, and siblings-in-law, when applicable.

"The Development or Growth of the Self" section was dedicated to assessing the individual's life from before conception through adolescence. With regards to prenatal life, this section asked the individual to consider the physical and mental conditions of their parents, their occupations during time of pregnancy, and any "unusual experiences" that occurred during pregnancy.²⁸ The other subsections of "The Development or Growth of the Self," which addressed infancy, childhood, and adolescence, asked similar questions about physical and mental development, with the two latter subsections incorporating questions on possible vocations, social

²¹ Allen, "The Eugenics," 239.

²² Wildman, "Psychoses of the Feeble-minded," 530.

²³ Wildman, "Psychoses of the Feeble-minded," 531.

²⁴ Allen, "The Eugenics," 238.

²⁵ Yerkes and LaRue, *Outline of a Study*.

²⁶ Yerkes and LaRue, *Outline of a Study*, 7.

²⁷ Yerkes and LaRue, *Outline of a Study*, 8.

²⁸ Yerkes and LaRue, *Outline of a Study*, 9.

tendencies, and special habits. They also addressed environmental influences on the individual such as diet, alcohol, and school experiences.²⁹

"The Self of Today" section continued to ask questions similar to the previous section, though with more detail and with reference to the individual in the present day. It asked the individual to consider, for example, how their personality related to their occupation or how prepared they were for marriage and parenthood.³⁰ Furthermore, this section included the questionnaire, "Index to the Germ-Plasm," germ-plasm referring to one's genetic makeup, in order for the individual to gauge the fitness of any potential offspring with their partner, if applicable.³¹ The survey asked questions similar to those of the "Record of Family Traits" questionnaire; however, "Index to the Germ-Plasm" more specifically addressed personality and mental traits. One of the prompts invited the individual to select various activities for which they had predilections, such as studying, reading, traveling, exploring, and dancing.³²

With the goal of demonstrating how their data was synthesized, the ERO published *The Family-History Book*, compiled by Davenport, in their seventh bulletin.³³ Though the data in the book was taken from ERO questionnaires and interviews, any real names of people were censored for confidentiality purposes. The book provided summaries of families with regards to their mental and physical traits. A few of the sections exhibited racial overtones, such as Section V, "Records of Recent Immigrants" and Section VI, "Records of Negro-White Crosses." Section V specifically examined what was deemed to be a desirable family from Sicily, Italy, providing an example of what type of family would supposedly match the social standards of the United States.³⁴ This case study commented on many of this family's qualities such as occupation, economic success, eye color, complexion, life span of deceased family members, and personality of the parents, children, and grandparents. Many of these family members had fair complexions and blue or brown eyes, and their father was deemed to have been economically successful. Furthermore, the author of this study commented on family members' personalities using positive terms such as energetic, honest, and easygoing. Unlike most sections in *The Family-History Book*, Section VI, "Records of Negro-White Crosses," did not provide a detailed description of the family featured in the case study. Instead, this section listed the hair colors and textures of individuals in a biracial family from Jamaica as well as the proportions of family members with the skin tones the author denoted as black, red, yellow, and white.³⁵ There were no indications of whether a certain proportion of traits regarding skin tone was advantageous or more fit for society.

Davenport did not explicitly state that certain racial or ethnic groups were inferior or superior to others, though race influenced much of the categorization systems of the Eugenics Record Office.³⁶ A definite influence was presented in *The*

²⁹ Yerkes and LaRue, *Outline of a Study*, 14.

³⁰ Yerkes and LaRue, *Outline of a Study*, 21.

³¹ *Ibid.*

³² *Ibid.*

³³ Charles Davenport, comp., *The Family-History Book* (New York: Cold Spring Harbor, 1912).

³⁴ Davenport, *The Family-History*, 46-53.

³⁵ Davenport, *The Family-History*, 52-53.

³⁶ Ifekwunigwe, 'Mixed Race', 102.

Trait Book, in which field workers were instructed to record when all members of a family presented characteristics such as “Negroid” or “Blond.”³⁷ Additionally *The Family-History Book*, compiled by Davenport, contained an entire section regarding race crossing.³⁸ Whenever race or place of origin was mentioned in one of Davenport’s ERO documents, the identifier was not given a positive nor negative connotation. This contrasts with the mental categorization system presented in Wildman Jr.’s *Journal of Nervous and Mental Disease*, in which “mongolian” was listed as a characteristic for an “idiot.” In turn, rather than condemning specific races as inferior, Davenport advocated against the mixing of different racial groups.

4. The Biological Foundation of Eugenics

Charles Davenport was the first eugenicist to apply Mendel’s Laws to human beings.³⁹ He believed that the majority of human traits were controlled by independent genes; therefore, he argued that the most important issue regarding race mixing was the resulting variability in traits. For example, he claimed that a person may have a gene relating to long legs and another relating to short arms, rather than a gene affecting the lengths of their limbs as a whole.⁴⁰ He would refer to such a person with disproportionately sized limbs as the result of a “disharmonious” cross because they were most likely born of a parent with long legs and another with short arms.⁴¹ Today, arguments against race mixing are often associated with the possible dilution of traits. The dilution argument against race mixing implies that a long-limbed and short-limbed person would produce offspring with medium-length limbs; however, Davenport’s argument implied that the offspring would have a greater variability, or what he regarded as a disharmony, in traits.

Although Davenport did not explicitly claim that certain ethnic groups were inferior to others, he advocated against the production of offspring between certain races. He wrote “Race Crossing in Jamaica” in 1929, in which he defined a racial group as a “group of individuals constituting a subdivision of the species characterized by the possession of some one distinctive hereditary trait.”⁴² He claimed that the categorization of races required an element of isolation so that their distinctive traits may develop; he drew attention to what he regarded as the societal issue of different racial groups having children with each other. He argued that the mixing of races that differed too greatly from each other would result in a considerable amount of variability of traits in the offspring and heighten the chances of the offspring possessing a disharmonious selection of traits.⁴³ In this work, he further elaborated on the consequences of race crossing, using a study he conducted in Jamaica as evidence.

Davenport utilized tests of mental ability to compare individuals belonging to categories of Black, Brown or Mulatto, and White. In this study, Brown and Mulatto

³⁷ Davenport, *The Trait*, 12.

³⁸ Davenport, *The Family-History*, 46.

³⁹ Ifekwunigwe, ‘*Mixed Race*’, 101.

⁴⁰ Ifekwunigwe, ‘*Mixed Race*’, 105.

⁴¹ Ibid.

⁴² Charles B. Davenport, “Race Crossing in Jamaica,” *The Scientific Monthly*, September 1928, 225.

⁴³ Ifekwunigwe, ‘*Mixed Race*’, 106.

were used as terms to describe biracial individuals with Black and White ancestry. The ways of examining the persons included musicality and rhythm tests, arithmetic math tests, and common-sense tests, among others. The Seashore test measured individuals' "ability to make fine discriminations in the elements of musical capacity."⁴⁴ The test was subdivided into pitch and rhythm and required examinees to distinguish between various patterns of each musical element. With regards to the pitch test, Black individuals, on average, scored 75 points, while White individuals, on average, scored 71 points. Brown individuals scored the highest, with an average of 77 points; however, Davenport found that there was greater variability in the scores of Brown individuals as opposed to the other two groups.⁴⁵

Davenport also implemented the Army Alpha tests, which included eight separate examinations of intelligence. One of the tests was centered around the ability to solve basic arithmetic problems. Again, Black examinees scored higher on average than their White counterparts, though Brown examinees scored in between the two groups on average.⁴⁶

Overall, Davenport concluded that White examinees had the best scores, though in some cases, Black and Brown examinees scored higher. Davenport claimed that for many of these tests, including the Seashore test and the Army Alpha tests, the scores of Brown people varied the most.⁴⁷ This was used as evidence for his argument that the mixing of different racial groups would lead to greater variability, and hence, greater disharmony in traits. He believed that the unpredictability of their characteristics posed a problem as well, in that mixed race offspring could either exhibit the greatest number of high-quality or low-quality traits in a given population.⁴⁸

In general, Davenport advocated against race crossing, even though he admitted that some biracial crosses would generate high-quality offspring. In the "Negro-White Crosses" section of the *Family-History Book*, he referred to some biracial individuals as possessing "admirable concentrations of traits."⁴⁹ Because he believed the range of traits being produced through the cross breeding of races was so unpredictable, however, he argued that society should not risk the mass production of low-quality offspring for the sake of saving a few high-quality offspring.⁵⁰

Table B of the "Race Crossing in Jamaica" study provided a summary for the mean scores received by each racial group in the eight Army Alpha tests. Scores in the group labeled "Black" ranged from 5.9 points to 15.8 points, with an average of 9.64. "Brown" scores ranged from 5.1 points to 12.7 points, with an average of 7.95 points. "White" scores presented the most variability, ranging from 4.9 points to 20.3 points, with an average of 10.23 points.⁵¹ Davenport consistently advocated against the mixing of different racial and ethnic groups because of the unpredictability of their offspring's genetic makeup; however, according to this data that he collected, the group with the greatest variety in intellect was the White category.

⁴⁴ Davenport, "Race Crossing," 234.

⁴⁵ Ibid.

⁴⁶ Davenport, "Race Crossing," 235.

⁴⁷ Davenport, "Race Crossing," 236.

⁴⁸ Ifekwunigwe, *Mixed Race*, 106.

⁴⁹ Davenport, *The Family-History*, 52-53.

⁵⁰ Ifekwunigwe, *Mixed Race*, 106.

⁵¹ Davenport, "Race Crossing," 236.

TABLE B
SUMMARY OF MEAN SCORES OBTAINED IN THE EIGHT ARMY ALPHA TESTS

	I	II	III	IV	V	VI	VII	VIII	Avg.
Black	5.9	10.0	5.9	15.8	8.8	7.2	13.9	9.6	9.64
Brown	5.1	8.4	5.2	12.7	6.4	5.6	10.8	9.4	7.95
White	4.9	7.5	8.5	20.3	11.4	6.8	10.2	12.2	10.23

Figure 1. Summary of the Army Alpha tests' mean scores according to racial group (Davenport, 1928)

5. Eugenics in the Greater Progressive Era

5.1 Critiques of Eugenics

The eugenics movement collapsed post-World War II due to anti-racism efforts. The vilification of the Nazi Party, which engaged in eugenic practices explicitly motivated by anti-Semitism, led to a vilification of eugenics itself.⁵² Hence, the eugenics movement lost a great deal of traction as the world took a step towards supporting egalitarianism. For instance, the United Nations Educational, Scientific, and Cultural Organization, UNESCO, created the Commission on Human Rights in 1946, advocating against discrimination based on identifiers such as race, sex, and religion.⁵³ There were critics of eugenics before World War II, however. Not only did these critics exist, but many of them were racist themselves. Many critics of eugenics believed in the field's overall goals of improving humanity in some way, only desiring to criticize eugenic organizations' methodology and scientific evidence rather than their values.

Dr. David Heron, a member of the Galton Laboratory for National Eugenics in London, criticized Davenport and the Eugenics Record Office for their methodology in data collection after watching Davenport's presentation at the International Eugenics Congress in 1912.⁵⁴ Heron's main criticisms centered around Davenport's use of Mendel's Laws of Inheritance to explain the inheritability of human traits. In the *New York Times* article, "English Expert Attacks American Eugenic Work," Heron commented on Davenport's beliefs that what he regarded as mental defectiveness was a recessive trait and that those considered mentally weak may produce offspring with those considered mentally strong.⁵⁵ Heron explained that if mental defectiveness were a recessive trait, the offspring of these proposed parents would have a fifty percent chance of being mentally defective. In that case, Heron believed that segregating the "mental defects" from the rest of society would result in lower numbers of undesirable offspring than allowing them to marry "normal" people.⁵⁶ Heron also criticized the use of Mendel's Laws of Inheritance as a whole, believing that the cause of what was regarded as feeble-mindedness was more

⁵² Deborah Barrett and Charles Kurzman, "Globalizing Social Movement Theory: The Case of Eugenics," *Theory and Society* 33, no. 5 (October 2004): 512.

⁵³ Barrett and Kurzman, "Globalizing Social," 512.

⁵⁴ David Heron, "English Expert Attacks American Eugenic Work," *The New York Times* (New York), November 9, 1913.

⁵⁵ Heron, "English Expert."

⁵⁶ *Ibid.*

complicated than the presence of a single recessive gene. Instead, he argued that the cause was an absence of one or more genetic factors in an individual. Despite criticizing the Eugenics Record Office, Heron was a eugenicist himself who worked for a eugenics lab in London.

Similarly, William Ernest Castle criticized Davenport's claims that every genetic trait was independent. Castle specifically drew attention to Davenport's idea that separate genes affected the various limbs of the human body; Castle argued that a singular gene affected the entire skeleton, differentiating between "general size factors" as opposed to "specific size factors."⁵⁷ Castle sought to disprove the general belief of eugenicists that an increased production of mixed-race offspring would lead to supposed racial deterioration. Through inspecting data on individuals of African, European, and indigenous descent and multiracial people, he concluded that a mixing of traits merely led to an "intermediate degree of the characters involved," rather than a great variability in high-quality and low-quality traits.⁵⁸ Regardless, Castle still believed that Black people overall were less intelligent than their White counterparts and that biracial people with a Black and White ancestry would exhibit less intelligence than White people. In the present day, rather than Davenport's variability model, this dilution model is more closely associated with the production of biracial or multiracial children. In other words, the way in which we think about race today is influenced by criticisms of eugenics in the early twentieth century.

5.2 The Progressive Era

The eugenics movement informed other movements throughout the Progressive Era. With the overlap between the women's suffrage and birth control movements and the eugenics movement, women were encouraged to take control over their bodies and involve themselves in more societal responsibilities.⁵⁹ One of these responsibilities was supposedly improving the human race. Many female activists used eugenics to advance their own agendas. Eugenic feminism involved women regulating their sexuality in order to choose the ideal partner and produce what would be deemed high-quality offspring.⁶⁰

Progressive Era White feminist literature often contained the common theme of American nativism. Charlotte Perkins Gilman, a prominent feminist and writer at the time, for example, supported the notion that African Americans and immigrants would contaminate what she believed to be the purity of the Anglo-Saxon race.⁶¹ She emphasized the importance of White women engaging in positive eugenics, or the act of increasing high-quality traits in a population. In addition, in her work, "A Suggestion on the Negro Problem," she proposed the idea of segregating African Americans deemed "below a certain grade of citizenship" in an enclosed region of every county or town of the United States.⁶² She explained that the goal of this segregation system was to aid certain African Americans in social evolution to allow them to become "self-supporting" like the White members of society. She proposed

⁵⁷ Ifekwunigwe, 'Mixed Race', 105.

⁵⁸ Ifekwunigwe, 'Mixed Race', 106.

⁵⁹ Seitler, "Unnatural Selection," 66.

⁶⁰ Ibid.

⁶¹ Seitler, "Unnatural Selection," 67.

⁶² Charlotte Perkins Gilman, "A Suggestion on the Negro Problem," *American Journal of Sociology* 14, no. 1 (July 1908): 80.

that individuals old enough to work would engage in agricultural labor or labor in mills and small shops. Children, on the other hand, would receive an education to “guarantee the fullest development possible to each [child].”⁶³ Gilman infused her writing with eugenic ideology in order to fuel her agenda of elevating White women in society. Her strategy for uplifting these White women entailed dehumanizing Black people, using eugenic principles such as the difference between high- and low-quality traits to justify her beliefs.

Margaret Sanger, a leading birth control activist in the early twentieth century, promoted similar ideas regarding ridding the United States of people with presumed unfit traits. Like Gilman, she used the eugenics movement to advance feminist reform, specifically focusing on birth control activism. Sanger opened the nation’s first legal birth control clinic, the Clinical Research Bureau, in 1923 in New York.⁶⁴ The clinic was even supported by eugenic organizations such as the ERO. Sanger appealed to those who supported eugenics by advertising birth control as a means of eliminating racial threat.⁶⁵ She appealed to women by advertising how they could govern their own fertility, ensuring that they collectively work to produce, as Sanger said, “more children from the fit, less children from the unfit.”⁶⁶

Moreover, the Progressive Era left a lasting impact on American society’s economy. With the departure from laissez-faire capitalism, in which the government possessed little to no control over the free market, economists looked for ways to regulate the market more concretely.⁶⁷ The establishment of a minimum wage, which led to increased unemployment, was particularly pervaded by eugenic thought. Frank Taussig, an American economist at the time, suggested two types of unemployable workers. One type was the disabled or aged, while the other type was the feeble-minded.⁶⁸ The term feeble-minded was often used by eugenicists such as Davenport in the Eugenics Record Office. Eugenic thinking was utilized as a justification for ridding the workforce of individuals deemed low-quality employees.

Though eugenics faced criticism in the early twentieth century, the field had the support of other reform movements throughout the Progressive Era. By examining how eugenics influenced these movements, the lines between what we today regard as social reform and what we condemn as horrendous behavior become blurred.

6. Conclusion

The Progressive Era is associated with activism and social reform, including the passing of labor laws, initiatives in public health and hygiene, the women’s suffrage movement, and the beginning of the birth control movement.⁶⁹ Eugenics not only coexisted with, but also helped fuel these various other movements with the same

⁶³ Gilman, "A Suggestion," 81.

⁶⁴ Richard A. Soloway, "The 'Perfect Contraceptive': Eugenics and Birth Control Research in Britain and America in the Interwar Years," *Journal of Contemporary History* 30, no. 4 (October 1995): 639.

⁶⁵ Soloway, "The 'Perfect,'" 642.

⁶⁶ Seitler, "Unnatural Selection," 67.

⁶⁷ "Retrospectives: Eugenics," 217.

⁶⁸ "Retrospectives: Eugenics," 214.

⁶⁹ Ballard Campbell, "Comparative Perspectives on the Gilded Age and Progressive Era," *The Journal of the Gilded Age and Progressive Era* 1, no. 2 (April 2002): 166-167.

overarching goal: to improve human society. Labor reformists such as Frank Taussig utilized eugenic principles to further the idea that some people are more deserving of employment over others. Birth control activists such as Margaret Sanger promoted the idea that women can take control of their bodies to produce fit offspring for society. The eugenics movement as a whole provided an opportunity for scientists to apply their theories to actual people.

Race heavily informed the ideologies and practices within the eugenics movement and the Progressive Era. Much of the reform brought about by the Progressive Era was limited to upper-class White Americans. Moreover, the eugenics movement was intended to bring about reform that would specifically benefit White Americans with the advocacy against race mixing and the production of low-quality traits, the criteria for which were sometimes racialized themselves. It may seem like common sense that scientific conclusions should be shaped on the basis of evidence; however, because of the racism embedded into eugenic thinking, evidence justifying eugenic practices seemed to be shaped on the basis of preset conclusions. This was seen in Davenport's "Race Crossing in Jamaica" study, in which his data did not correspond to his overall claim that White people generally possess the most predictable, and therefore most desirable, sets of traits.⁷⁰

Despite the racism that was ever-present throughout the eugenics movement, even critics of eugenics did not argue against the values of the field, merely desiring to find fault with the methodology and use of scientific facts. Anti-racist efforts only rose to prominence post-World War II, leading to the collapse of the eugenics movement. The criticisms of eugenics prior to this total collapse are an essential piece in understanding how the movement coexisted with other methods of social reform throughout the Progressive Era. The ERO ceased operation on December 31, 1939, marking a transition period from "old-style" eugenics to "new-style" eugenics.⁷¹ Eugenic reform was pioneered by Frederick Osborn, one of the founding members of the American Eugenics Society.⁷² Whereas forced sterilization as a means of executing eugenic thinking was the main focus of old-style eugenics, new-style eugenics involved population control through genetic counseling and medical genetics, only subtly promoting sterilization.⁷³

Regardless, today the eugenics movement is given a negative connotation while the rest of the Progressive Era is given a positive one. We enjoy the successes of the Progressive Era and celebrate the period's activists for pioneering social reform. Who are we truly celebrating, though? To what extent do we ignore several movements' ties to eugenic thought? Perhaps we need to reconsider the lens through which we view the Progressive Era, taking into account the nuances of the ideologies of activists we praise today.

Bibliography

- Allen, Garland E. "The Eugenics Record Office at Cold Spring Harbor, 1910-1940: An Essay in Institutional History." *Osiris* 2 (1986): 226-64.
- . "The Misuse of Biological Hierarchies: The American Eugenics Movement, 1900-1940." *History and Philosophy of the Life Sciences* 5, no. 2 (1983): 105-28.

⁷⁰ Davenport, "Race Crossing," 236.

⁷¹ Allen, "The Eugenics," 254.

⁷² Stern, *Eugenic Nation*, 115.

⁷³ *Ibid.*

- Barrett, Deborah, and Charles Kurzman. "Globalizing Social Movement Theory: The Case of Eugenics." *Theory and Society* 33, no. 5 (October 2004): 487-527.
- Blount, Anna E. "Eugenics in Relation to Birth Control." *The Birth Control Review*, January 1918, 7.
- Campbell, Ballard. "Comparative Perspectives on the Gilded Age and Progressive Era." *The Journal of the Gilded Age and Progressive Era* 1, no. 2 (April 2002): 154-78.
- Carr-Saunders, Alexander Morris. "A Criticism of Eugenics." *The Eugenics Review*, October 1913, 214-33.
- Davenport, Charles, comp. *The Family-History Book*. New York: Cold Spring Harbor, 1912
- . *The Trait Book*. N.p.: Lord Baltimore Press, 1912.
- Davenport, Charles B. "Race Crossing in Jamaica." *The Scientific Monthly*, September 1928, 225-38.
- Davenport, Charles Benedict. *Heredity in Relation to Eugenics*. N.p.: Henry Holt and Company, 1915.
- Galton, Francis. *Inquiries into Human Faculty*. New York: Macmillan, 1883.
- Gilman, Charlotte Perkins. "A Suggestion on the Negro Problem." *American Journal of Sociology* 14, no. 1 (July 1908): 78-85.
- Hemmenway, Henry B. Review of *Eugenics Record Office, Report No. 1 (1913)*. *Journal of the American Institute of Criminal Law and Criminology* 5, no. 4 (1914): 623-26.
- Heron, David. "English Expert Attacks American Eugenic Work." *The New York Times* (New York), November 9, 1913.
- Ifekwunigwe, Jayne O., ed. *'Mixed Race' Studies: A Reader*. N.p.: Taylor & Francis Group, 2004.
- Laughlin, Harry H. *Report No. 1*. N.p.: Eugenics Record Office, 1913.
- . *Report of the Committee to Study and to Report on the Best Practical Means of Cutting off the Defective Germ-plasm in the American Population: The Scope of the Committee'S Work*. N.p., 1914.
- "Retrospectives: Eugenics and Economics in the Progressive Era." *The Journal of Economic Perspectives* 19, no. 4 (Fall 2005): 207-24.
- Seitler, Dana. "Unnatural Selection: Mothers, Eugenic Feminism, and Charlotte Perkins Gilman's Regeneration Narratives." *American Quarterly* 55, no. 1 (March 2003): 61-88.
- Soloway, Richard A. "The 'Perfect Contraceptive': Eugenics and Birth Control Research in Britain and America in the Interwar Years." *Journal of Contemporary History* 30, no. 4 (October 1995): 637-64.
- Stern, Alexandra Minna. *Eugenic Nation: Faults and Frontiers of Better Breeding in Modern America*. Second ed. N.p.: University of California Press, 2016.
- Wallin, J. E. Wallace. *The Education of Handicapped Children*. N.p.: Houghton Mifflin Company, 1924.
- Wildman, H. Valentine, Jr. "Psychoses of the Feeble-minded." *The Journal of Nervous and Mental Disease* 42, no. 8 (1915): 529-39.
- Yerkes, Robert M., and Daniel W. LaRue. *Outline of a Study of the Self*. N.p., 1913.



Film's Portrayal of Mental Illness and Violence: Schizophrenia, DID, and PTSD

Brian Nguyen

Author Background: *Brian Nguyen grew up in the United States and currently attends Northside College Preparatory High School in Chicago, Illinois in the United States. His Pioneer research concentration was in the field of media studies and titled "Media and Society."*

Abstract

Film is a medium that has a large impact on the general public. In films, mental illness is a popular theme that is often depicted in an insensitive and inaccurate manner. Specifically, schizophrenia, dissociative identity disorder (DID), and post-traumatic stress disorder (PTSD) are some of the most commonly depicted mental illnesses in film. These three illnesses also happen to be often depicted in films along with violence, which creates and reinforces stereotypes against people with the illnesses. This stigmatization of mental illness harms the communities of people who suffer from symptoms of the mental illness on a daily basis. Though many films' portrayal of mental illness perpetuates negative stereotypes and misconceptions of mental illness, there are accurate portrayals of mental illness in films that allow the audience to better understand the illness.

1. Introduction

From the onset of cinema, mental illness has been a topic many films aim to explore through their characters and stories. Specifically, violence and mental illness are often linked with one another, usually as recurring themes throughout many different films. The stereotype started with mad killers and mad doctors, which were popular in early cinema (Wahl, 1995, p. 56). An example of this is the film *The Cabinet of Dr. Caligari* from 1919, which is about a mad, evil doctor who uses his knowledge for nefarious reasons. This trend of turning mentally ill characters into villainous murderers and criminals continued throughout the twentieth century. These characters, largely depicted as villains in films like *Psycho* and more recently *Joker*, are portrayed as people who are unhinged and mentally unstable due to their mental illnesses, which drive them to commit heinous acts of violence for no other reason than them

being crazy. According to a study done by Stuart (2006), one in four mentally ill fictional characters (in media as a whole) kills someone and half hurt others. Furthermore, the offense rate of mentally ill characters with speaking parts is 30%, while other television characters only have an offense rate of 3%. This study shows how mentally ill characters tend to be more villainized, as they are usually depicted as being more violent. In addition to the insane killer trope, violence through war and self-harm is also often seen linked with mental illness in films. Whether the violence be projected on the individual or the people around them, mental illness is sometimes associated with it.

Mental illnesses, as defined by the American Psychiatric Association, are “health conditions involving changes in emotion, thinking or behavior (or a combination of these)” (Parekh, 2018). The stigmatization of these different mental illnesses harms people who suffer from these illnesses on a daily basis. Of the mental illnesses that have been adapted to the silver screen, schizophrenia, dissociative identity disorder (DID), and post-traumatic stress disorder (PTSD) are some of the most commonly represented disorders, leading to the public having varied understandings of them.

However, many films handle mental illness in a sensitive and informative manner, so the audience gains a better understanding of mental illness. Despite the potential for film to be a medium that can help educate the general public on issues concerning mental health, the stigma against mental illness has still largely been reinforced throughout history by film (Wahl, 1995), especially in regard to schizophrenia, DID and PTSD.

2. Schizophrenia

Schizophrenia has been a center of focus in many films. According to the Mayo Clinic, “Schizophrenia is a serious mental disorder in which people interpret reality abnormally. Schizophrenia may result in some combination of hallucinations, delusions, and extremely disordered thinking and behavior that impairs daily functioning, and can be disabling” (Mayo Clinic, 2020). Psychotic symptoms of schizophrenia include delusions, hallucinations, and thought disorder, while cognitive symptoms include difficulty concentrating and processing information (NIMH, n.d.). People with schizophrenia usually start exhibiting symptoms during early adulthood, anywhere from the range of late teens to mid-thirties (NIMH, n.d.). There is no current cure for schizophrenia, though research is being conducted to find better ways to treat the illness in patients. Schizophrenia only affects about one percent of the world’s population (NIMH, n.d.).

In a study from 2012, Owen examined the portrayal of schizophrenia in different films. Owen analyzed 72 movies produced from 1990 to 2010. In total, 42 characters met selection criteria for schizophrenia. The study showed that a majority of the characters, 83%, displayed dangerous or violent behaviors toward others (Owen, 2012). Additionally, 31% of the characters that were violent engaged in homicidal behavior (Owen, 2012). However, not all movies depicting schizophrenia are bad portrayals. In fact, one of the earliest depictions of schizophrenia in film is *The Snake Pit*, the 1948 film directed by Anatole Litvak,

which follows Virginia Cunningham, a mentally unstable woman with schizophrenia who is in a mental hospital struggling to recall the origins of her mental illness (Tett, 2006). The film is often praised for being ahead of its time (Zimmerman, 2003). Its intent was to raise awareness of the poor treatment of mentally ill patients.

A modern example of schizophrenia's depiction in film is *Shutter Island*, the 2010 film directed by Martin Scorsese, which follows U.S. Marshal Teddy Daniels as he investigates the disappearance of a mentally ill murderer from an asylum on *Shutter Island* (Cox, 2010). The twist revealed at the end of the movie is that Daniels is not a real person, only a delusion of patient Andrew Laeddis, who is implied to have schizophrenia (Cox, 2010). Mental illness, used in this way, feels like a mere plot device to shock the audience with an unexpected, cheap twist. Furthermore, there are many violent moments in the film. In a study DeMare conducted to analyze the frequency of violence in schizophrenic protagonists in different films, findings show that Daniels has nine instances of violence committed against others throughout the movie (DeMare, 2016). The violence, in this case, is easily linked back to Daniels's mental illness, as the film paints him in a negative light as a crazed man. This is detrimental to people who suffer from schizophrenia, as it provokes the audience to associate schizophrenia with violence, especially after viewing the character commit nine acts of violence throughout the movie. They may assume that because of the twist of mental illness, that is the root of the violence in the character. Though the movie received mostly praise with some mixed reviews, conflating mental illness with violence only helps to further stigmatize schizophrenia among the general public. Its popularity proves that films can continue stigmatizing mental illness while still being profitable with critical acclaim, which is a dangerous precedent to set. Additionally, another modern example of schizophrenia portrayed in film is *A Beautiful Mind*, a 2001 film directed by Ron Howard, which is often cited as being, for the most part, accurate in its depiction of schizophrenia in the main character (Alyssa, 2021). The movie follows mathematician John Nash in the 1950s as he struggles with schizophrenia while still managing to make significant contributions to the field of mathematics, eventually going on to win a Nobel Prize (Bauer, n.d.). The film's depiction of schizophrenia makes the character empathetic, humanizing him in a way that allows the audience to understand his emotions and experiences, "specifically the realistic struggle with medication adherence in this patient population" (Thacker and Hughes). However, there are critiques by some mental health experts of the exaggerated symptoms that Nash experiences in the film. In the movie, Nash has many visual hallucinations, though in reality, hallucinations are more likely to be auditory than visual (Alyssa, 2021). Despite this fact, Howard opted to make the story more cinematic to keep the film visually engaging. Even then, *A Beautiful Mind* is one of the most well-known films centered around schizophrenia, notable for its mature handling of the illness. Because the film is realistic in its portrayal of schizophrenia, it can be useful for educational purposes (Rosenstock, 2003). The manner in which the illness is shown makes it a good film for patients, families, and the general public to learn about schizophrenia (Rosenstock, 2003).

3. Dissociative Identity Disorder

Dissociative identity disorder, similarly to schizophrenia, is a mental illness that is often misunderstood by the general public. Common misunderstandings are that DID is not real or that it is the same as schizophrenia (Peisley, 2017). According to the American Psychiatric Association, “Dissociative identity disorder is associated with overwhelming experiences, traumatic events and/or abuse that occurred in childhood. Dissociative identity disorder was previously referred to as multiple personality disorder,” with the main symptom being “the existence of two or more distinct identities (or ‘personality states’). The distinct identities are accompanied by changes in behavior, memory and thinking. The signs and symptoms may be observed by others or reported by the individual” (Wang, 2018). As stated in the definition, DID is usually sparked as a reaction to trauma, often during childhood, as a way to cope with the traumatic memories. Because of this, people who experience physical and/or sexual abuse as children have increased risks of developing DID (Wang, 2018). The inherent nature of DID makes it easy to dramatize in films, as the presence of multiple personalities in a character is used many times as a driving force of conflict in plots.

Early depictions of DID in film can be traced back to movies like *The Three Faces of Eve*, where a woman named Eve White discovers through the help of a psychiatrist that she has DID, with three personalities in total. Her psychiatrist further helps her rediscover the childhood trauma that she experienced that caused her to develop DID (Duffin, 1996). While the film is critically acclaimed and helped to bring attention to the disorder among the general public and mainstream audiences, it fails to show the actual complexities of DID, instead opting to naively depict the illness and psychiatry (Perring, 2006). Despite being slightly outdated in its execution, it serves as a dramatic and informative film, considering it came out in 1957.

A modern example of DID’s depiction in film can be found in *Split*, the 2016 movie directed by M. Night Shyamalan. In the film, the main character Kevin Wendell Crumb has DID, with 23 different personalities existing in him throughout the film. Though most of the alters are not dangerous, there is one named “The Beast” that is violent and maniacal, leading to the abduction of three teenage girls (Dodgson, 2019). The film paints Kevin as a madman, a dangerous person being forced to kidnap girls because of his mental illness. Mental health professionals argue that the film stigmatizes DID, “raising the potential for dangerous attitudes to emerge and for people with the illness to be damaged,” according to Elizabeth Howell, a psychotherapist from New York (Healthline, n.d.). Kevin, in the movie, is antagonized and depicted as a crazed man with a killer and dangerous personality hiding underneath the surface. This depiction demonizes actual people who suffer from DID, most of whom do not have dangerous and violent alters (Healthline, n.d.). In a study conducted on 173 individuals with DID, researchers found that of the sample they researched, “3% reported having been charged with an offense, 1.8% were fined, and 0.6% were incarcerated in the past 6 months. No convictions or probations in the prior 6 months were reported” (Webermann and Brand, 2017). This study shows how individuals with DID are not inherently more likely to be violent, so the problematic antagonization of Kevin in the film serves to perpetuate an

oversimplification that people with DID are violent criminals. In a study done by Sampson (2020), eight films from 1995-2020 that portray at least one character with DID were analyzed. The method was split into four sections. The first was testing the characters for DID through DSM-5 criteria. The second was analyzing the violent and criminal behavior in the eight different films by counting the instances of violent behavior and instances of criminal behavior. The third was examining the characters for extreme alternative identities and instances of incompetence. The fourth was counting the instances of stigmatizing behavior or disparagement through stigmatizing words/terms (Sampson, 2020). The results were that the movies had “commonly depicted exaggerated stereotypes and misconceptions about this disorder” (Sampson, 2020, pg. 86). One of the movies examined was *Split*, which depicted 8 instances of violent behavior and 5 instances of criminal behavior, as well as instances of incompetence (Sampson, 2020). This study further shows how the film repeatedly portrays violent acts, associating them with Kevin, which in turn pushes the audience to link DID with violence and blame the disorder for the cause of that violence. Whether the movie meant to or not, it vilifies people with DID, which only furthers the misconception that people with mental illnesses are violent beings.

4. Post-Traumatic Stress Disorder

Lastly, post-traumatic stress disorder is another mental illness often depicted in film. According to the American Psychiatric Association, “PTSD is a psychiatric disorder that may occur in people who have experienced or witnessed a traumatic event such as a natural disaster, a serious accident, a terrorist act, war/combat, or rape or who have been threatened with death, sexual violence or serious injury” (Torres, 2020). Though it is commonly associated with war veterans, about 3.6% of adults in the US report experiences with PTSD every year, with approximately 1 in 11 people estimated to be diagnosed within their lifetime (Torres, 2020). PTSD is not limited to only war trauma, as it requires exposure to an upsetting traumatic event, either firsthand or indirectly. Therefore, anyone can develop PTSD at any age (Torres, 2020). Symptoms of PTSD include intrusive thoughts and/or flashbacks of the traumatic event, avoidance of places or objects that remind them of the event, and cognitive changes in the form of negative thoughts and emotions as well as self-destructive behavior (Torres, 2020). People with PTSD are more likely to be violent towards themselves than towards others. In “Anger, Aggression, and Self-Harm in PTSD and Complex PTSD,” researchers conducted a study on 44 participants with different backgrounds and ages who have experienced a “Troubles-related incident” and examined them for PTSD and self-harm (Dyer, Dorahy, Hamilton, Corry, Shannon, MacSherry, McRobert, Elder, McElhill, 2009). They found that 95% of them met diagnostic criteria for PTSD. Sixty-six percent of them had a history of self-harm. Thirty-nine percent had had a history of self-harm only and twenty-seven percent had a history of both self-harm and attempted suicide.

As PTSD is a common disorder that affects combat veterans, it is many times associated with wartime violence. Movies like *American Sniper* and *The Deer Hunter* portray the violence that soldiers endure through times of war,

whether it be American soldiers during the Vietnam War or the more recent Iraq War. An early and iconic film depicting PTSD is *The Deer Hunter*. In *The Deer Hunter*, the story follows three main characters, Nick Chevotarevich, Michael Vronsky, and Steven Pushkov, all of whom are Vietnam war veterans. The film shows the viewer the characters from before they go to war, during their time in the war, and after the war (Bauer, n.d.). This structure gives insight into the behavior of the soldiers when they led normal lives and their lives subsequently after the traumatic events that took place during the war. The big catalyst of the PTSD is the three soldiers' capture by the Viet Cong, where they are imprisoned and forced to play Russian Roulette for the guards' amusement (Bauer, n.d.). Russian Roulette is a deadly game where the player places a single round in the chamber of a revolver, spins the cylinder, and pulls the trigger of the gun while pointing it at themselves. The three are able to kill their captors and escape. Back in America, they find it difficult to reassimilate back into their old lives, as their friends and family do not understand them or their experiences in the war (Bauer, n.d.). The contrast between the three protagonists' lives before and after the war is stark, as they grapple with readjusting to their normal lives after living through such a traumatic event (Bauer, n.d.). The alienation that the three feel is common among war veterans with PTSD (Torres, 2020). The film's portrayal of PTSD is accurate and harrowing, similar to veterans' experiences in real life. PTSD, in the case of this film, is not as problematic in its portrayal as that of *Split* or *Shutter Island*, but the conflation of PTSD and wartime violence reinforces the association between mental illness and violence.

A modern example of PTSD's depiction in film can be found in *American Sniper*, the 2014 film directed by Clint Eastwood. Based loosely on the autobiography of Chris Kyle, sniper during the Iraq War, the film takes creative liberty in its portrayal of Kyle on his tours of battle during the war. In the movie, Kyle goes on four total tours in Iraq, working as a sniper helping his fellow soldiers (Clyman, 2015). He goes on to kill many people, including women and children, who were threatening the lives of American troops, and witnesses the deaths of his close friends (Clyman, 2015). Because of his experiences in the war, he develops PTSD in between tours, which are the times in which he is at home with his family. He shows many of the common symptoms of PTSD, such as being lost in thought and anxiety from being away from the battlefield (Clyman, 2015). When he is at home, he is constantly on edge and easily panicked by sounds, like a baby crying (Lee and Straw, n.d.). His symptoms only worsen with each tour, and in the end, he is able to ease his symptoms through helping other war veterans, slowly recovering to normalcy himself (Lee and Straw, n.d.). The movie is modified from the book to create a more cinematic film and dramatic story, thus it is not an entirely autobiographical film (Lee and Straw, n.d.). Despite the changes to better fit Hollywood, the film still overall does a good job in portraying the PTSD and trauma that soldiers can go through (Lee and Straw, n.d.).

5. Impact of Stigmatization of Mental Illness in Film

The stigmatization of mental illness in America is widespread, as a 2002 poll discovered that 82% of people agree that there is a stigma attached to mental

illness, with 66% (among parents) saying they would be uncomfortable if a person with a serious mental illness worked in their child's school (Roper Center for Public Opinion Research, 2015). Movies like *Split* can cause people to believe in misconceptions of mental illness, which furthers the stigma of mental illness (Sampson, 2020). In the case of mental illness, because mental illness and violence have been linked in films for so long, the general public over time has come to think of the mentally ill as dangerous and violent individuals, backed by the films and media they consume.

In the study, "Impact of the Film, 'One Flew Over the Cuckoo's Nest,' on Attitudes Towards Mental Illness," George Domino and the researchers aimed to explain the impact of films that depict mental illness on the public's views of mental illness. The experiment featured 146 college students who were all given a questionnaire on their attitudes towards mental illness (Domino, 1983). Then, after *One Flew Over the Cuckoo's Nest* was released, they were given the same questionnaire again, with 85 people having seen the film and 39 having not seen it. Finally, they were given it again after half of them watched a documentary intended to balance the film's portrayal of life in a mental institution. They found that there were no significant differences between the students before the release of the film, but after it came out, people who saw the movie were less positive about mentally ill people (Domino, 1983). The documentary did nothing to help their attitudes (Domino, 1983). The results in this study are surprising, due to the fact that *One Flew Over the Cuckoo's Nest* was actually largely a critique of psychiatry and the inhumane ways that society treated the mentally ill in the 1950s. Despite the intended message of the film, based on the experiment, audiences still left the movie with overall more negative attitudes towards mentally ill people than beforehand, with or without the documentary that was intended to serve as a disclaimer that the film was fiction (Domino, 1983).

Many films featuring violent and mentally ill antagonists risk viewers automatically associating the two as conjoined traits. Even if a film is created for pure entertainment and not educational purposes, there is still risk that the stereotypes from the movie affect the way the viewer perceives mental illness. For example, in the 2019 film *Joker*, there are themes of both mental illness and violence growing gradually in the character of the Joker throughout the movie until the end (Alexander, 2019). The implications of this are that violence and mental illness go hand in hand, and that the reason for the Joker's violent side coming to light is because of his mental illness worsening. Pozios argues that this should not be true, as his violence is not linked to any delusions or hallucinations and stems from revenge (Alexander, 2019). Still, audiences may come to the conclusion that because the film spends so much time focusing on the Joker's symptoms of mental illness, which build up to him embracing violence, this means that the reason that the Joker becomes a violent murderer is because of his untreated mental illness (Alexander, 2019). This conclusion is dangerous to those that suffer similar symptoms to the Joker, as they may feel ostracized and insecure that people may perceive them as violent and dangerous. Though about 10% of people who are mentally ill are violent, 90% are neither dangerous nor violent, so this relation is a lot more complicated than how it is portrayed on the screen (Zimmerman, 2003).

Beyond film, other sources of media also provide content centered around mental illness. Specifically, YouTube provides a platform for people with

DID to explain and show their mental illness on the internet in the form of videos, with the goal of helping viewers gain a better understanding of the illness (Feidelson, 2021). The YouTube DID community is large, with some channels, like MultiplicityAndMe, having over 200,000 subscribers (Feidelson, 2021). However, since dramatic portrayals of DID garner more views and followers than realistic and tame ones, creators may be tempted to sensationalize their mental illness (Feidelson, 2021). In addition to impacting public view of DID, this also has the potential to worsen dissociative symptoms, since the creators are “intentionally cultivating the differences between the parts for an online audience” (Feidelson, 2021).

6. Accurate Portrayals of Mental Illness in Film

Plenty of films and TV shows that portray mental illness in a sensitive and educational manner exist. *Mr. Robot*, a show created by Sam Esmail, features Elliot Alderson as the main character, a cybersecurity engineer by day, vigilante hacker by night, who suffers from DID and depression (Watson, 2016). Despite the show depicting both violence and mental illness throughout all four seasons, it received widespread acclaim from both audiences and medical professionals (Watson, 2016). One reason for its praise is its portrayal of the different alters that Elliot has (Watson, 2016). While films like *Split* have all the personalities played by one actor, *Mr. Robot* has different actors for different personalities. This helps to make the characters more distinct and feel like they are actually different people, which is what people with DID experience and relate to (Watson, 2016). In *Split*, since James McAvoy is playing all of the personalities, to the audience it can seem absurd and abnormal, further driving home the stigma that DID and mental illness is not normal. In *Mr. Robot*, however, since the story is from the point of view of Elliot, the audience views the characters as separate because to him, they are separate people. This depiction makes Elliot feel more human and sane, because the audience is viewing everything from his mind (Watson, 2016). Not only does this allow the viewer a better understanding of how people with DID operate and think, it also is good storytelling and usage of mentally ill characters without being exploitative. However, *Mr. Robot* is still criticized by some for being sensationalized; Feidelson wrote that *Mr. Robot* is among many media where DID is “represented as alternately freakish and alluring” (Feidelson, 2021). Because the show depicts Elliot’s alters as different people, some argue that they are perpetuating the myth that the different personalities are actually different people as opposed to stemming from one person’s imagination (Wang, 2018). Ultimately, the character of Elliot Alderson, despite suffering from several mental illnesses, is a relatable and empathetic person that makes the audience root for him despite his shortcomings.

Conversely, *Canvas*, a film directed by Joe Greco, is about Mary Marino, a mother and wife suffering from schizophrenia, as her husband Joe and 10-year-old son Chris try to adjust to living with her while learning to make sense of the illness. The story is told largely from the point of view of Chris, as the film, despite being fictional, was inspired by director Greco’s childhood growing up with a mother with schizophrenia (Cunningham, 2007). Gabbard wrote that “Greco had

a mission in making this film: 'I want people to know that mental illness is not a death sentence. There's hope'" (Gabbard, 2007). Mary is faced with both delusions and visions that she does not understand, yet that does not detract from her love for her family (Cunningham, 2007). The film does a great job of humanizing Mary in a way that shows the audience how horrific dealing with schizophrenia really is for the person suffering from it (Cunningham, 2007). In addition to Mary's realistic reactions to her schizophrenia, her husband and son's reactions to her illness are also similar to how family members in real life behave in their situation (Cunningham, 2007). Joe quits his job to begin work on a sailboat in hopes of teaching his family how to sail, avoiding the typical tropes of a Hollywood ending when Mary is unable to join them (Cunningham, 2007). The film accurately shows how there is no cure for schizophrenia, and that even with treatment, symptoms still worsen over time. This serves to show the audience how serious and heart-wrenching the illness is, as no matter how much they try, the family still feels helpless at the hands of a mental illness they barely understand (Cunningham, 2007). The film has been praised by many mental health professionals as being realistic and accurate to the real-life illness, including Glen Gabbard M.D., who also states that "While the film is heartbreaking, it leaves the audience with a sense that schizophrenia is something that can be endured, managed, and assimilated by a family intent on surviving it" (Cunningham, 2007).

Lastly, in the case of PTSD, the 2008 film *The Hurt Locker*, directed by Kathryn Bigelow, is about Staff Sergeant William James and his bomb defusal squadron as they deal with warfare in Iraq, set during the Iraq War. The movie is critically acclaimed for its realism in both depictions of war and trauma. In the film, James is depicted as a reckless man when it comes to putting himself and his squadron in danger, though despite this, he is extremely gifted in defusing bombs (Bauer, n.d.). Throughout the film, James and his squadron are faced with dangerous and violent moments in the middle of the war zone, being forced to act quickly on their toes. Though James's recklessness leads to several deaths and injuries, his gift at defusing bombs proves to save many people. By the end of his rotation, he makes it out of Iraq alive and arrives home to his wife and son (Bauer, n.d.). However, he finds that he cannot acclimate back to his normal life, and decides to go back to Iraq to continue fighting. The movie was written by Mark Boal, a journalist who took inspiration from his own personal experiences embedded in a bomb squad during the Iraq War (Bauer, n.d.). At the end of the film, James is so affected by his experiences in the war that he cannot even do mundane tasks like grocery shopping, as he is too distraught from his PTSD (Serlin, 2010). The movie's theme of war being a drug is echoed with his choice to go back to Iraq, showing how his trauma affected him so deeply that he feels he cannot readjust to living his previous life, and would rather go back to face the terrors of the battlefield where he feels most comfortable.

All of the films listed so far have been created with the primary purpose of entertainment. In addition to fictional portrayals of mental illness in films, documentaries can be extremely useful in terms of educational purposes for the general public in learning more about the realities of different mental illnesses. As many fictional films are made with the intent of entertaining rather than educating, documentaries can serve as a good use of the film medium to teach audiences about mental illnesses that are often misrepresented in an engaging

and entertaining way. In the article “An Exploration of How Film Portrays Psychopathology: The Animated Documentary Film *Waltz with Bashir*, The Depiction of PTSD And Cultural Perceptions,” Hankir and Agius analyze the film *Waltz with Bashir*, an animated autobiographical film about a soldier living with PTSD from the Sabra and Shatila massacre in Beirut. He lost his memory of the event, and in the end he is able to regain the memories with the help of a psychologist, and it ends with footage of actual victims from the massacre (Hankir and Agius, 2012). The film as a whole is an anti-war movie, and it gives us an insight into a soldier’s struggles with PTSD (Hankir and Agius, 2012). Hankir and Agius argue that this insight is good for teaching people about mental illness and psychopathology, which in turn allows for the public to empathize more with people struggling with mental health (Hankir and Agius, 2012). Though the film is a documentary, it is unique in its storytelling, opting for animation to depict the story in a narrative style.

7. Conclusion

Portrayals of mental illness in films throughout history have been for the most part problematic, whether inadvertently or purposefully. Though there are exceptions, violence is often conflated with mental illness, with mental illness being blamed as the cause of violence in antagonists such as murderers and criminals. Due to this, the general public often misunderstands these illnesses and makes assumptions that the mentally ill are dangerous and violent, conclusions which are usually drawn from the media they consume. Schizophrenia, DID, and PTSD are all popular mental illnesses in the media, particularly in film.

Despite the stigmas that films perpetuate, film can still be a good medium for informing the public about mental illness in an educational manner. Past examples show that good stories involving mental illness do not have to be sensationalized or about violence. In fact, films that humanize people who suffer from mental illness can be more interesting than those that serve only to villainize people with mental illness. The best films and TV shows that feature mental illness have characters that the audience can empathize with. If the viewer can understand the struggle of the mentally ill character, the film is successful in both creating a good story and destigmatizing mental illness. Films and shows like *Mr. Robot*, *Canvas*, and *The Hurt Locker* prove that there is much potential for future media to raise awareness of mental health issues while still creating thought-provoking and captivating pieces. As entertainment media is a source of information for many people, it is important that the content being produced and consumed is not harming and exploiting communities of people. Future research should be done on the effects of film on the stigmatization of mental illness on the general public, especially in the context of modern films.

Works Cited

- Alyssa. (2021, May 10). How schizophrenia was portrayed in a beautiful mind. Mental Health Program at Banyan Treatment Centers. <https://www.banyanmentalhealth.com/2021/03/30/how-schizophrenia-was-portrayed-in-a-beautiful-mind/>
- Bauer, P. (n.d.). A Beautiful Mind. *Encyclopedia Britannica*. <https://www.britannica.com/topic/A-Beautiful-Mind>
- Bauer, P. (n.d.). The Deer Hunter. *Encyclopedia Britannica*. <https://www.britannica.com/topic/The-Deer-Hunter>
- Bauer, P. (n.d.). The Hurt Locker. *Encyclopedia Britannica*. <https://www.britannica.com/topic/The-Hurt-Locker>
- Clyman, J. (2015, January 19). American Sniper. *Psychology Today*. <https://www.psychologytoday.com/us/blog/reel-therapy/201501/american-sniper-0>.
- Cox, D. (2010, July 29). Shutter island's ENDING explained. *The Guardian*. <https://www.theguardian.com/film/filmblog/2010/jul/29/shutter-island-ending>.
- Cunningham, R. (2007, July 2). Canvas: A Movie with Many Dimensions. *Health Central*. <https://www.healthcentral.com/article/canvas-a-movie-with-many-dimensions>.
- DeMare, N. (2016). "Exaggerations and Stereotypes of Schizophrenia in Contemporary Films." *Elon Journal of Undergraduate Research in Communications*, 7(1). Retrieved from <http://www.inquiriesjournal.com/a?id=1474>
- Dodgson, L. (2019, January 10). Dissociative identity disorder is nothing like how it's portrayed IN 'split,' according to people who have it. *Business Insider*. <https://www.businessinsider.com/living-with-split-personality-disorder-2017-2>.
- Domino, G. (1983). Impact of the Film, "One Flew over the Cuckoo's Nest," on Attitudes towards Mental Illness. *Psychological Reports*, 53(1), 179–182. <https://doi.org/10.2466/pr0.1983.53.1.179>
- Duffin, J. (1996, October 8). The three faces of eve. NYU Langone Health. <https://medhum.med.nyu.edu/view/10022>.
- Durham, R., & Wilkinson, P. (2020). Joker: How 'entertaining' films may affect public attitudes towards mental illness – psychiatry in movies. *The British Journal of Psychiatry*, 216(6), 307-307. doi:10.1192/bjp.2020.56
- Dyer, K. F., Dorahy, M. J., Hamilton, G., Corry, M., Shannon, M., MacSherry, A., McRobert, G., Elder, R., & McElhill, B. (2009). Anger, aggression, and self-harm in PTSD and complex PTSD. *Journal of Clinical Psychology*, 65(10), 1099–1114. <https://doi.org/10.1002/jclp.20619>
- Feidelson, L. (2021, August 10). Meet My Multiple Mes People with dissociative identity disorder have started presenting their selves on YouTube to rapturous fans. *The Cut*. <https://www.thecut.com/2021/08/dissociative-identity-disorder-youtube.html>.

- Gabbard, G. (2007, October 5). Schizophrenia on Filmmaker's Canvas. *Psychiatric News*. https://psychnews.psychiatryonline.org/doi/10.1176/pn.42.19.0006?utm_source=TrendMD&utm_medium=cpc&utm_campaign=Psychiatric_news_TrendMD_0.
- Hankir, A., & Agius, M. (2012). An exploration of how film portrays psychopathology: the animated documentary film *Waltz with Bashir*, the depiction of PTSD and cultural perceptions. *Psychiatria Danubina*, 24 Suppl 1, S70–S76.
- Lee, S., & Straw, K. (n.d.). *Movies in the Classroom: American Sniper*. CPNP. <https://cpnp.org/perspective/2015/06/259881>.
- Mayo Foundation for Medical Education and Research. (2020, January 7). Schizophrenia. Mayo Clinic. Retrieved September 10, 2021, from <https://www.mayoclinic.org/diseases-conditions/schizophrenia/symptoms-causes/syc-20354443>.
- Owen, P. R. (2012). Portrayals of schizophrenia by entertainment media: A content analysis of contemporary movies. *Psychiatric Services*, 63(7), 655–659. <https://doi.org/10.1176/appi.ps.201100371>
- Parekh, R. (2018, August). What is Mental Illness? *American Psychiatric Association*. <https://www.psychiatry.org/patients-families/what-is-mental-illness>.
- Peisley, T. (2017, January 10). Busting the myths about dissociative identity disorder. Retrieved from <https://www.sane.org/information-stories/the-sane-blog/mythbusters/busting-the-myths-about-dissociative-identity-disorder>
- Perring, C. (2006, September 12). Review of "The Three faces of Eve". Review - The Three Faces of Eve - Dissociative Disorders. https://www.helenfarabee.org/poc/view_doc.php?type=book&id=3295&cn=41.
- Roper Center for Public Opinion Research. (2015, April 28). Public Attitudes about mental health. Retrieved from <https://ropercenter.cornell.edu/public-attitudes-about-mental-health>
- Rosenstock, J. (2003). Beyond a beautiful mind: Film choices for teaching schizophrenia. *Academic Psychiatry*, 27(2), 117–122. <https://doi.org/10.1176/appi.ap.27.2.117>
- Sampson, V. (2020). *The Portrayal of Dissociative Identity Disorder in Films*. Elon University.
- Serlin, I. A. (2010, March 7). Hurt locker: Treating trauma in the body and the ptsd experience. *Psychology Today*. <https://www.psychologytoday.com/us/blog/make-your-life-blessing/201003/hurt-locker-treating-trauma-in-the-body-and-the-ptsd-experience>.
- Stuart, H. (2006). Media portrayal of mental illness and its treatments: what effect does it have on people with mental illness? *CNS drugs*, 20(2), 99–106. <https://doi.org/10.2165/00023210-200620020-00002>
- Tett, Joan (2006). "The Snake Pit and Social Awareness and Treatment of Mental Illness," *ESSAI*: Vol. 4 , Article 37.
- Thacker, M., & Hughes, C. (2013). Movie review: A beautiful mind. *Mental Health Clinician*, 2(8), 246–247. <https://doi.org/10.9740/mhc.n132978>
- Torres, F. (2020, August). What is posttraumatic stress disorder? What Is PTSD? <https://www.psychiatry.org/patients-families/ptsd/what-is-ptsd>.

- U.S. Department of Health and Human Services (n.d.). Schizophrenia. National Institute of Mental Health. <https://www.nimh.nih.gov/health/topics/schizophrenia>.
- Vogel, D. L., Gentile, D. A., & Kaplan, S. A. (2008). The Influence of Television on Willingness to Seek Therapy. *Journal of Clinical Psychology*, 64(3), 276-295. DOI: 10.1002/jclp.2044
- Wahl, O. F. (1995). *Media madness: Public images of mental illness*. Rutgers University Press
- Wang, P. (2018, August). What Are Dissociative Disorders? American Psychiatric Association. <https://www.psychiatry.org/patients-families/dissociative-disorders/what-are-dissociative-disorders>.
- Watson, Z. (2016, August 3). How 'Mr. Robot' nailed its depiction of mental illness. Inverse. <https://www.inverse.com/article/19084-mr-robot-mental-illness-dissociative-identity-disorder-split>.
- Webermann, A. R., & Brand, B. L. (2017). Mental illness and violent behavior: the role of dissociation. *Borderline personality disorder and emotion dysregulation*, 4, 2. <https://doi.org/10.1186/s40479-017-0053-9>
- White, C. (2012, August 17). Distortions plague portrayals of schizophrenia on film. *Psychiatric News*. https://psychnews.psychiatryonline.org/doi/10.1176/pn.47.16.psychnews_47_16_7-a.
- Zimmerman, J. N. (2003). *People like ourselves: Portrayals of mental illness in the movies*. Lanham, Md: Scarecrow Press.



Making Decisions in a Crisis: The Ethics of Scarce Resource Distribution Policies During COVID-19

Caroline Greenstone

Author Background: *Caroline Greenstone grew up in the United States and currently attends Greenhill School in Addison, Texas in the United States. Her Pioneer research concentration was in the field of philosophy and titled “Bioethics.”*

1. Introduction

Since March 2020, the coronavirus (COVID-19) pandemic has shifted attention towards the healthcare infrastructure designated with the task of saving lives, but amid the apparent scarcity of resources, the question became: whose life gets to be saved? For so long, scientists and doctors warned about a pandemic of similar magnitude to the one before us today (Schmidt, 2020). The pandemic, which to the “average” individual may have seemed like a shock, was highly predictable. In America, the warning signs, available to the government before the public, did not instigate the necessary rapid preemptive response (Graham, 2020). In its place, as COVID began to devastate communities across the country, states made “Crisis Standards of Care” policies. For years, bioethicists and healthcare professionals have debated and written papers surrounding the very questions of how doctors should make decisions regarding the allocation of scarce resources. But distinct from the sometimes-abstract debates that have occurred in the past, policies now have tangible implications.

In this paper, I will argue that distributing scarce medical resources during crisis settings on the basis of judgments of the patient’s past or future quality of life, life expectancy, or worth in society is unethical because it violates principles of patient care.

1.1 The Status of COVID

The status of COVID is constantly changing. As of August 2021, there have been 628,000 deaths in America and 4.33 million deaths worldwide. The purpose of this paper is not to analyze specific policies currently in place in America or anywhere else, but rather to analyze decisions in crisis, determine the extent to which these decisions are ethical, and build on previous literature to propose a new way to decide who should receive scarce resources. The widespread lack of resources often forces decisions that many want to mitigate or delay by dubbing them “impossible situations.” Regardless of how we refer to them, their ethics are definitely not impossible to analyze.

1.2 Resource Scarcity

The idea of resource scarcity is quite common in bioethical research and discourse, being that scarcity is almost always a condition of healthcare systems across the world (Supady et al., 2021, pp. 430–434). Medical ethics discourse, focused on the issue of scarcity, became exceptionally prominent in the 1960s following the establishment of a decision-making committee at a Seattle hospital, often unofficially labeled the “God Committee.” A set of individuals gained the power to allocate the use of a scarce kidney machine without any guidelines. These individuals began to make decisions based on constituency and age. They then considered other factors: family life, income, gender, and societal worth (Glass and Sullivan, 2020). The nature of these decisions led to a degree of public outcry that set the tone for many bioethical debates, including the one that I am participating in with this paper.

Medical triage is the concept of determining which patients can be saved in a limited-resource, often crisis, setting. Triage, used by and credited to military operations, categorizes patients into groups based on severity, degrees of urgency, and whether available resources can be of use (Vinay et al., 2020, p. 6). Of these categories, one is always a level of severity that is unable to be solved under present situations, even if possible under normal circumstances—this practice labels efforts of revival as medically futile. The concept of medical futility is often attributed to Hippocrates, who wrote that individuals should not be treated if they were “overmastered by their diseases” (National Council on Disability, 2019, p. 21).

There are multiple ways communities and doctors go about making decisions regarding care. There is a spectrum on which these decisions exist. The National Academy of Medicine labeled this spectrum “A Continuum of Care.” The spectrum includes three different levels: conventional care (that which is received daily in typical settings), contingency care (where demand “begins to exceed supply”), and crisis care—the situation in which COVID has left the country in (National Academies of Sciences, Engineering, and Medicine, 2020, p. 4).

The University of California Critical Care Bioethics Working Group describes the situation in this way: “Crisis Standards of Care (CSC) are applied when a health system is so overwhelmed by a public health event that critically ill individuals, who would normally receive any reasonable therapy, may receive limited treatment or non-traditional provision of care if the system surge capacity is exceeded” (University of California Critical Care Bioethics Working Group, 2020, p. 6).

There are many places where a Crisis Standard of Care policy is in place. These policies take two different forms. One is implicit-based processes that have doctors make decisions on an individual basis. This approach is often thought to be individualized but is critiqued for lowering the patient quality of care and placing undue psychological burdens on doctors to make these decisions. The other is explicit, where there is a clear policy for how doctors should analyze a case and make a decision. The latter approach is often appreciated for providing a transparent mechanism for the decision process and promoting equality of decisions (Spector-Bagdady et al., 2020, pp. 77–80). Notably, however, even with explicit policies in place, doctors are often still the actors implementing the

guidelines given to them. This is to say that objectivity is still often lacking even within explicit policies because they are subject to the individual assessments of doctors when regarding the relevant factors.

Some explicit triage policies, as set out by Nicholas Rescher in his paper “The Allocation of Exotic Medical Lifesaving Therapy,” include at least two steps in determining who gets a resource: exclusion factors (often based on biomedical factors), and inclusion factors, which are far more vague, even conceivably subjective (such as potential worth and family value) (Rescher, 1969, pp. 173–186). Rescher’s paper was written following the decisions of the “God Committee,” but it was so influential that many suggested policies use similar ideas today. Rescher’s process for distributing what he labeled as “exotic medical lifesaving therapy” consisted of exclusion, inclusion, and a randomized process between the patients included. The basic model of exclusion factors and inclusion factors is fairly consistent among many policies. I am inclined to agree with the additional third step outlined by Vinay et al. (2021): repeated assessment of a patient’s condition. This is not to say that previous authors would not agree with repeated assessments as well. Rescher’s utilitarian perspective would probably mean that he would agree. However, my point is that repeated assessment deserves heightened importance in these discussions. Insofar as life-saving treatment is lengthy and chances of survival can change as treatment progresses, a part of the conversation must be responsive to the fluidity of the condition.

2. Ethics of Resource Scarcity

2.1 Ethical Principles Behind Clinical Care

Throughout history, our collective understanding of bioethics has shifted. At the center of most dialogue surrounding medical ethics is Hippocrates and the Hippocratic Oath. What was once seen as an impermeable set of rules and guidelines, to be gatekept by the professionals tasked with implementing them, has now “become a subject of widest public concern” (Pellegrino, 1993, p. 27). The Hippocratic Oath contained ideas still pertinent today: notions of beneficence, nonmaleficence, and confidentiality. But amid the rise of new technology, changing practices in patient interaction, and social justice movements, a new category was included: justice (Pellegrino, 1993, pp. 27–31). Beneficence refers to a doctor’s obligation to act in the patient’s best interest, while nonmaleficence regards a doctor’s obligation to do no harm. And justice, sometimes also associated with or referred to as “equality,” refers to the role of the doctor in providing fair distribution of resources and treatment to all. Today, all of these categories play a critical role in patient care.

2.2 Ethical Principles Behind Public Health

The goals of public health often prioritize equality in access and distribution, respect for individuals and patients, and addressing disparities in past treatment and care (Blacksher, 2012, pp. 320–331). However, public health policies often fail to accurately recognize the problem, placing the blame for disparities in health on individuals. Blacksher’s (2012) argument is that:

The policy and practice of targeting disadvantaged groups takes place within and is itself a part of a broader context of social relations and institutional processes that subordinate nondominant groups either by (1) heightening conceptions of them as deviant, thus denying their common humanity, or (2) imposing on them expectations of assimilation, thus denying their distinctive life experiences, traditions and histories. (pp. 322–323)

As will become apparent throughout this paper, COVID has yet again subjected individuals to the failures of public health institutions and society as a whole. Policies that prevent individuals from accessing critical resources during COVID because of comorbidities or underlying conditions further disenfranchise already marginalized communities and place the responsibility of health solely on the individual.

2.3 Ethics During COVID

The primary objective of any ethical standard or norm is to tell us how to operate within our lives. The problem arises when the very nature of our lives rapidly and drastically changes. This is not to say that ethics go out the window during a national crisis: quite the opposite. It is my firm belief that our ethics must be stronger than ever. Every patient cared for under COVID is a patient cared for under the context of a national public health crisis.

For this reason, the central guiding principles in determining policies must be beneficence, nonmaleficence, justice (which includes recognition and equality), respect, objectivity, and accountability (transparency).

While public health policies have often used a utilitarian calculus when allocating scarce resources in crisis settings (Haward et al., 2020, p. 38), too often, the guise of “maximizing utility” has been used as a façade for implementing policies that do not fairly distribute resources or provide equal respect and recognition. I do not believe it is fair to characterize utilitarianism as “saving the greatest number of lives.” This is because seeking to maximize utility is centered within the status quo and tends to diverge from conversations about future implications. To be clear, my argument is not that utilitarian approaches are invalid. My argument is that in order to promote ethical care, deontological and egalitarian policies are preferable and increase individual quality of care. Utilitarian policies, which by nature ignore how individual patients are treated if a long-term objective is accomplished, can become a source of discrimination (Supady et al., 2021, pp. 430–434). Responding to COVID is a matter of saving lives. I argue that in order to operate ethically, individual patients must be treated ethically, so improving patient care means centering the values of justice and equality, which also saves lives in the future because it sets an important precedent for how individuals should be treated. While amid crises like COVID, it can be easier to focus on the present, I do not think it is justifiable to ignore the future implications of our actions.

3. Is It Ethical to Allocate Scarce Resources on the Basis of Quality of Life, Life Expectancy, or Worth in Society?

While biomedical factors are often used to determine who is excluded from treatment or given access to scarce resources, in many places, vague and sometimes subjective measures are used to determine who is included. Although, this is not the case in every current setting, previous Crisis Standard of Care policies such as those in Washington, Alabama, Oregon, Tennessee, and many other states cast a light on the ethics of those decisions. Although some of the aforementioned states have changed policies since initially announcing them, the question I have posed seeks to determine their ethics, so we can better understand how to craft future policies and make difficult decisions. In this section, I will be analyzing the inclusion factors that are used when the decisions about allocation are not randomized.

3.1 Life Expectancy

It is frequently argued that treatment should go to those who would benefit from it the longest. This statement is often used as a pretext to claim that those who are older, or have lived longer lives, should be deprioritized when the only distinctive factor is age. For example, a hospital has only one ventilator left, and there are two patients who need it. Every biomedical factor is the same, which means both patients are equally in need of the ventilator. The patients have similar chances of surviving and benefiting from the treatment. Also, co-ventilation is not a possibility in this situation because of differing physiological needs. It is in this situation that those who make the previously stated argument will choose to give the scarce resource to the person who is younger. To clarify, there is a substantive difference between deciding who should get treatment because one will not survive and deciding who should get the treatment because of how long one survives after it. I do not take issue with the former.

There are two different types of standards related to life expectancy that exist within explicit policies: those which clearly demarcate age guidelines and those that just say that life expectancy can be considered. The first is problematic "at the margins" (Stone, 2020, p. 580). It inflicts a new degree of arbitrariness into the decision-making process, as individuals only days apart can be differentiated along a line that does not truly influence how long they would benefit from treatment. In doing so, rigid life expectancy guidelines violate important standards of justice.

The second standard is not as overtly problematic; however, in my eyes, it still qualifies as something to leave out of future assessments and decisions. There are three main problems that I have with this standard. First, life is inherently unpredictable. No matter a person's age, they can suddenly and unpredictably die. This is to say that a young, previously healthy patient could get hit by a bus on their way out of the hospital. And tragically, this could happen to anyone. While the general public may wish to the contrary, doctors do not have the power to know how long a person has left to live. While the projected lifespan of an individual may be available when making decisions, events occur that are totally out of the control of any individual. And while doctors may be able to

justify their decision in certain circumstances by citing the potential of life in younger individuals, that justification would probably not be applicable in other circumstances. For example, while one may be willing to decide that someone who is 17 (my age) should get a ventilator over someone who is 70, that decision may seem more “difficult” if the question is between someone 17 and 10, or even a ten-year-old versus a newborn. This makes decision-making on the basis of life expectancy unethical because objectivity is at the core of promoting equality and justice in a pandemic. And it is my opinion that in order to make objective decisions in crisis settings, the justification for using a certain norm must be universalizable.

The second apparent problem is the lack of equity within the life expectancy “tie-breaker.” As John Stone (2020) wrote, “past and continued injustices are strong reasons why potential life years saved should not influence triage priorities for individuals who prospectively will likely survive more than a few months after hospitalization” (p. 581). Comorbidities that could potentially influence lifespan are, in some instances, a result of historical injustices that targeted marginalized communities (Ettachfini, 2020).¹ To illustrate, environmental racism subjects many communities of color to dangerous air quality and unclean water which causes illnesses that decrease life expectancy. This is to say that life expectancy is inherently an inequitable model in determining whether someone should access critical resources.

I want to acknowledge what may seem like a contradiction to those who may be reading this paper. It is very true that such comorbidities can sometimes make it less likely someone would survive the treatment in the first place, meaning they would not be prioritized for it. While it would be desirable to resolve these issues or find a different and more ethical way to allocate scarce resources, I do not believe that we now have the tools to do that at our disposal. Perhaps crafting policies predicated on more egalitarian ideas will allow us to reshape public health in the future, creating more equitable policies going forward.

Third, there is an immense amount of bias towards the elderly exhibited by doctors and the rest of society (Stone, 2020, p. 581). When the standard is not set clearly, and I have already discussed why a clearly set standard is arbitrary and unethical, doctors often make the decisions themselves. Because of the implicit bias that many individuals exhibit, this decision has the potential to be extremely dangerous for the elderly.

3.2 Quality of Life

In March 2020, as COVID and the fear of this pandemic’s implications grew, states implemented guidelines for rationing healthcare and medical resources. In Washington State, a policy was put in place that advised hospitals to consider the “baseline functional status” of a patient (Romano and Bagenstos, 2020). Quickly, this policy drew backlash, as it should. Policies or any decisions that take in the functional status, physical or mental, of a patient are rooted in ableism. The National Council on Disability wrote, “Providers often perceive people with

¹ Ettachfini’s point was those past injustices, targeting communities of color, have led to an increase of diseases and worsening health conditions for those communities. I added that such comorbidities have the potential to decrease life expectancy.

disabilities to have a low quality of life when, in reality, most report a high quality of life and level of happiness, especially when they have access to the healthcare services and supports that they need to equally participate in and contribute to their communities” (National Council on Disability, 2019, p. 10). The attempt to extrapolate one’s “functional status” and determine their quality of life is not only a failure in patient care but sets a poor precedent for the future. It is impossible for individuals to fully comprehend the lived experiences of others. Using quality of life assessments is nothing more than setting a value judgment on the quality of life with a disability. Humans have equal moral worth (Stone, 2020, p. 579). Any policy that seeks to mitigate or dismiss this truth is unjust, causes undue harm to patients, and strips individuals of their autonomy.

When decisions to provide treatment are rooted in a subjective judgment about the quality of one’s life with preexisting conditions, it worsens the quality of care for individuals who are already subject to lower standards of support, with minimal representation and advocacy within healthcare professions (Romano and Bagenstos, 2020). It prevents individuals from accessing care when they have an equal right to treatment. This constitutes an injustice, and by unfairly preventing access to care, these decisions directly contribute to an inferior standard of care for people with disabilities. If a doctor has an ethical obligation to “do no harm,” that doctor is actively avoiding their duty to a patient when using quality of life assessments.

Additionally, assessments that regard a patient’s quality of life after treatment also have the potential to be problematic. These decisions directly contribute to the system-wide biases against people with disabilities and pre-existing conditions, insofar as it is a doctor implicitly saying that a life is not worth saving because of the potential for lasting conditions. However, I do not believe that this specific form of assessment is absolutely and wholly unethical. Rather, I recognize that there are medical factors that can be quantified, and there are potentially severe repercussions to treatment. Nonetheless, doctors should seek to avoid making such decisions because of the potential for long-lasting harm.

Finally, if a patient is unable to express their own wishes or assessment of their quality of life, then a doctor is unable to morally justify not treating that patient as an equal because the doctor can never understand the complexity of the patient’s lived experiences. In that case, a doctor must not make assumptions about a patient’s desires. In response to arguments such as mine, many people point out examples where a family member or doctor must decide what to do when a patient has suddenly become severely injured or ill. Those types of examples are distinguishable because of the abrupt onset of the condition and the rapid nature of the decision-making. While COVID is a crisis setting and necessitates fast-paced decisions, it is distinct because the decision to put someone on a ventilator does not tend to occur within as limited a time. If a patient cannot speak for themselves, quality of life judgments are ill-informed and strip patients of their right to self-determination, in the present and in the future, by virtue of the decision.

3.3 Worth in Society

In “The Allocation of Exotic Medical Lifesaving Therapy,” Nicholas Rescher set out “acceptable” standards for final selection criteria for distributing resources:

likelihood of success, life expectancy, family role, prospective service, and retrospective service (Rescher, 1969, pp. 173–186). Rescher argued that allocating scarce resources is an “investment” that society is entitled to receive a “return” upon.

Rescher wrote, “In its allocation of ELT, society ‘invests’ a scarce resource in one person as against another and is thus entitled to look to the probable prospective ‘return’ on its investment” (Rescher, 1969, p. 178).

I think this is a particularly abhorrent norm for determining who accesses care. No individual should be objectified or viewed as a tool for society. Rescher sought to defend his point by arguing that although applying this standard is difficult, that does not mean we should not apply it (Rescher, 1969, p. 178). The matter of determining someone’s social utility is not something that is solely “difficult;” it requires one to clearly delineate their value system in a way that sets individuals and their “capabilities” against one another. I presume that in the vast majority of situations where a resource is being allocated, a doctor does not have full knowledge about the potential or future of their patients beyond medical knowledge, and I also presume that in most situations, the social utility of an individual is not going to be so clear. This means quite a few things. First, a doctor’s individual understanding of societal worth is going to pave the way for discrimination and unjust decisions. Society has a repugnant track record with using social utility in various forms. Second, doctors face a far greater psychological burden making these decisions because they consider factors beyond medicine, which lowers patient care for a greater number of people. Third, individuals are further degraded in society. Decisions made with individual patients are reflected in the way people interact with each other outside of the hospital. If social utility becomes a widely accepted standard, then it risks the further degradation of individuals.

I am aware that some of the arguments made above may be conceived as being a pragmatic implementation critique for a question of principle. In many circumstances, this may be a flawed method of argumentation. However, in this circumstance, it is necessary to view our principles through the lens of how they will be or have been implemented.

As Paul Ramsey wrote, “the equal right of every human being to live, and not relative personal or social worth, should be the ruling principle” (Ramsey, 1970). It is for this reason that randomization is a necessity, even while acknowledging that there are many differences in the ways that individuals conduct their lives or just exist. Even further, my argument is not that it is impossible to determine social utility or worth in all situations. Rather, it is that we should not do it, especially in times of triage or crisis-setting care.

Consistent with this view, I believe providing resources to healthcare workers must be a priority. I believe in this scenario, healthcare workers fall under the category of what Ramsey called: “narrowly defined, exceptional purposes” (Ramsey, 1970). While this prioritization may initially be seen as a contradiction, there is a distinction between evaluating one’s instrumental value within a specified setting and one’s moral value. As Emanuel et al. (2020) wrote, “These workers should be given priority not because they are somehow more worthy, but because of their instrumental value: they are essential to pandemic response” (p. 2053).

4. Potential Long-Term Implications

The implications of maintaining and creating policies that are unethical and using the standards listed above are immense, not just for current patients but for the future. The COVID pandemic is not over by any means. Hospitals across the country are once again being overwhelmed by these ostensibly impossible situations and decisions. Throughout the pandemic, rationing continues beyond ventilators (Iserson, 2020, p.482). There is rationing in personal protective equipment for doctors, nurses, and many other hospital workers, and there is rationing in vaccines and vaccine distribution across the world. It is far more difficult to operate under these conditions without an explicit policy in place. While this pandemic is devastating as it is, unethical decisions make it exponentially worse. In order to save lives and in order to respect life, our focus should just as equally be on the ethics of our decisions as on solutions to the pandemic. In my eyes, they are close to being one and the same.

5. Suggestion for New Crisis Standard of Care Policies

The Crisis Standard of Care policy that I recommend falls along similar lines to that which was put into place in the University of California Health system. First, there must be an objective exclusion step on the basis of medical factors. This step should ask the questions: “Is the patient’s condition severe enough to necessitate treatment, wherein this resource is the only option?” and “Is the patient likely to survive and benefit from treatment?” This step can be done through some form of calculated assessment, wherein patients’ needs and conditions are grouped, and treatment is prioritized via group (I am inclined to completely recommend the allocation assessment published by the University of California, insofar as it regards only medical considerations and whether the individual is a critical worker). The second step, when determining who among the group will receive treatment, should be randomized. This can be done through a “lottery system” but should not be done through a “first-come, first-served” system. The third step is far more continuous, requiring repeated re-allocation assessments. These assessments should measure the status of the patient’s condition and reanalyze the patient in the context of others, who may require greater medical care. This is only justifiable when all other options have been exhausted.

One question that has not yet been dealt with in this paper is the question of who is making this decision. Up until this point, I have operated under the assumption that it is doctors and the patient’s care team distributing the resources. This has been the assumption because it is what appears to have happened and is happening in many scenarios. Although policies (such as that of the University of California) implied that a “triage officer/team” would be put into place, I don’t believe that this is entirely feasible. Supday et al. (2020) wrote, “Experience during the early COVID-19 pandemic suggests that these committees would not be as flexible, nimble, and comprehensive enough to be responsible for all bedside decisions as they arise 24 [hours] a day and throughout the hospital, including emergency department, acute care wards, traditional ICUs, and makeshift ICUs” (p. 433). If a triage team were able to consult on every case and objectively make these decisions, I would support that. However, I do not believe

this is practical. It is for this reason that I believe policies must be explicit, objective, and just.

Last, this policy/norm should not just apply to patients suffering from COVID but should rather be extended to anyone in the hospital needing the same resource.

6. Addressing Counterarguments

In defense of the methods/norms critiqued in this paper, many claim the following: the difference in social utility is so clear that it can sometimes warrant usage in determining the allocation of care, and quality of life assessments are valid in situations outside of COVID, so they can be justified now as well. In this section, I will refute both.

The underlying assumption in many arguments that defend social utility, as in discussions of the “God Committee” referenced previously in the paper, is that it can be conflated with moral worth. It absolutely cannot. One’s independent, substantial, and equal worth is not the same as a person’s ability to contribute or function positively within the boundaries of what society deems as being “good.” The idea of social utility is simply a calculation of how well one can acquiesce to societal norms. I think the historical pretext within this statement is quite obvious. Social utility has constantly been used by oppressors to justify oppression. However, I will grant there are some scenarios in which social utility may seem clear. For example, the choice for allocating a scarce resource is between a doctor who provides free and accessible healthcare to children and an adult going to a high school to sell and pressure kids to buy/take drugs. Many would say the former is far more valuable in society and has far greater worth. I wouldn’t say that I disagree. However, my objection to using this measurement is not because it can’t be used or because it can never be clear. Rather, it is twofold. The crisis setting means not all the facts will be readily available. And second, we cannot determine social utility in many situations. Although it may be ostensibly easier to choose between a doctor and a drug dealer, it becomes far more complicated, even impossible, when the question is between a businessman and a janitor, for example. If we cannot use it for everyone, we cannot use it at all, because it takes away the objectivity and equality necessary in an ethics-oriented decision-making process.

On quality-of-life assessments, it is misplaced to argue that because something is a norm during “regular times,” it is justifiable during a crisis. There are two main reasons for this. First, crises are unique in that time is extremely limited and all of the information is not easily available. Second, just because these assessments are used outside of COVID and are a norm does not prove that they are ethical. Many argue that family members regularly make the decision of whether or not to stop care for their loved ones. However, not only are family members not the ones making this assessment in distribution circumstances, but time and information are scarce resources as well.

7. Conclusion

I believe that crises such as the one before us today serve as critical junctures for society, not solely because they are devastating and affect so many spheres of influence, but also because they challenge the way that we, as a society, perceive and operate in regard to our ethics. If we choose to stand down and acquiesce to the idea that these questions are “impossible” or “too difficult to answer,” we are choosing to let future generations be tarnished by ethical standards and norms that are subpar, to say the least.

In this paper, I first defined the ethics of certain resource allocation standards during a national public health crisis, and then I argued that quality of life judgments, life expectancy, and societal worth are all unethical standards for determining who should receive care.

By violating norms like beneficence, justice, equality, and respect, these standards corrode ethics. Crisis Standard of Care policies should be put into place and alleviate the pressure on doctors by giving clear and transparent guidelines. It may be necessary to change these policies based on the needs of a state. Still, the general idea should remain constant: determine the patient’s medical needs and randomize allocation within those groups. Updates to Crisis Standard of Care policies by several states are a good start, but as new COVID variants begin to overwhelm communities, all policies need to be reanalyzed and quickly readjusted.

Last, this paper does not intend to critique the many doctors, nurses, or hospitals who have been pushed to the breaking point throughout the pandemic and have worked extraordinarily hard to save as many lives as they could. This paper is not an attack on their work. Rather, it is for the millions of individuals currently in a hospital, who are desperate for treatment but are unable to access it. The decisions made are undeniably difficult and will be painful to some regardless of the factors that go into making them. However, we must be able to live with and justify these decisions so they do not continue to affect society years in the future. The impacts of unethical decisions now will extend beyond our lifetimes. In a time overwhelmed by grief, we cannot let our ethics and morals become yet another loss.

Works Cited

- Blacksher, E. (2012). Redistribution and Recognition [PDF]. *Cambridge Quarterly of Healthcare Ethics*, 21(3), 320-331. <https://doi.org/10.1017/S0963180112000047>
- Callahan D. (1988). Aging and the ends of medicine. *Annals of the New York Academy of Sciences*, 530, 125–132. <https://doi.org/10.1111/j.1749-6632.1988.tb35304.x>
- Chen, B., & McNamara, D. M. (2020). Disability Discrimination, Medical Rationing and Covid-19. *Asian Bioethics Review*, 12, 511-518. <https://doi.org/10.1007/s41649-020-00147-x>

- Emanuel, E. J., Persand, G., Upshur, R., Thorne, B., Parker, M., Glickman, A., Zhang, C., Boyle, C., Smith, M., & Phillips, J. P. (2020). Fair Allocation of Scarce Medical Resources in the Time of Covid-19 [PDF]. *The New England Journal of Medicine*, 382(21).
- Ettachfini, L. (2020, April 14). Coronavirus Death Rates Are a Direct Result of Environmental Racism. *Vice*. Retrieved August 26, 2021, from <https://www.vice.com/en/article/k7ev93/coronavirus-death-rates-environmental-racism>
- Glass, I. (Host), & Sullivan, L. (Narr.). (2020). Making the Cut (No. 729) [Audio podcast episode]. In L. Misitzis, E. Baker, & K. Mingle (Producer), *This American Life*. This American Life. <https://www.thisamericanlife.org/729/transcript>
- Graham, D. (2020, April 29). Why Trump Was Deaf to All the Warnings He Received. *The Atlantic*. Retrieved August 27, 2021, from <https://www.theatlantic.com/ideas/archive/2020/04/how-many-warnings-did-trump-ignore/610846/>
- Haward, M. F., Janvier, A., Moore, G. P., Laventhal, N., Fry, J. T., & Lantos, J. (2020). Should Extremely Premature Babies Get Ventilators During the COVID-19 Crisis? *The American Journal of Bioethics: AJOB*, 20(7), 37–43. <https://doi.org/10.1080/15265161.2020.1764134>
- Hodge, J. G., Jr, Hanfling, D., Hick, J. L., & Piatt, J. L. (2021). Diminishing disparities in U.S. crisis standards of care: Medical and legal challenges. *The Lancet*, 34. <https://doi.org/10.1016/j.eclinm.2021.100838>
- Iserson K. V. (2020). Healthcare Ethics During a Pandemic. *The Western Journal of Emergency Medicine*, 21(3), 477–483. <https://doi.org/10.5811/westjem.2020.4.47549>
- Laventhal, N., Basak, R., Dell, M. L., Diekema, D., Elster, N., Geis, G., Mercurio, M., Opel, D., Shalowitz, D., Statter, M., & Macauley, R. (2020). The Ethics of Creating a Resource Allocation Strategy During the COVID-19 Pandemic. *Pediatrics*, 146(1), e20201243. <https://doi.org/10.1542/peds.2020-1243>
- Loefler I. (2002). Why the Hippocratic ideals are dead. *BMJ: British Medical Journal*, 324(7351), 1463.
- National Academies of Sciences, Engineering, and Medicine. (2020, March 28). *Rapid Expert Consultation on Crisis Standards of Care for the COVID-19 Pandemic*. The National Academies Press. <https://doi.org/10.17226/25765>.
- National Council on Disability. (2019, November 6). *Quality-Adjusted Life Years and the Devaluation of Life with Disability* (Bioethics and Disability). National Council on Disability. https://ncd.gov/sites/default/files/NCD_Quality_Adjusted_Life_Report_508.pdf
- National Council on Disability. (2019, November 20). *Medical Futility and Disability Bias* (National Council on Disability, Author; Bioethics and Disability). https://ncd.gov/sites/default/files/NCD_Medical_Futility_Report_508.pdf
- Pearson, S. D., Hall, J. B., & Parker, W. F. (2020). Two for one with split- or co-ventilation at the peak of the COVID-19 tsunami: Is there any role for communal care when the resources for personalised medicine are exhausted?. *Thorax*, 75(6), 444–445. <https://doi.org/10.1136/thoraxjnl-2020-214929>
- Pellegrino, E. (1993). The Metamorphosis of Medical Ethics [PDF]. *JAMA*, 269(9).
- Persad G. C. (2020). A Conceptual Framework for Clearer Ethical Discussions About COVID-19 Response. *The American Journal of Bioethics : AJOB*, 20(7), 98–101. <https://doi.org/10.1080/15265161.2020.1779400>

- Ramachandran, P., Swamy, L., Kaul, V., & Agrawal, A. (2020). A National Strategy for Ventilator and ICU Resource Allocation During the Coronavirus Disease 2019 Pandemic. *Chest*, 158(3), 887–889. <https://doi.org/10.1016/j.chest.2020.04.050>
- Ramsey, P. (1970). A Human Lottery? In P. Ramsey (Author), *The Patient as Person: Explorations in Medical Ethics* (pp. 305-308) [PDF]. Yale University Press.
- Ranney, M. L., Griffeth, V., & Jha, A. K. (2020). Critical Supply Shortages — The Need for Ventilators and Personal Protective Equipment during the Covid-19 Pandemic [PDF]. *The New England Journal of Medicine*, 382(18). <https://doi.org/10.1056/NEJMp2006141>
- Rescher, N. (1969, April). *The Allocation of Exotic Medical Lifesaving Therapy*. The University of Chicago Press Journals.
- Romano, N., & Bagenstos, S. (2020, April 6). Don't deny ventilators to disabled patients. *Washington Post*. <https://www.washingtonpost.com/outlook/2020/04/06/coronavirus-ventilators-disabled-people/>
- Schmidt, C., & Undark. (2020, June 13). Coronavirus Researchers Tried to Warn Us. *The Atlantic*. <https://www.theatlantic.com/health/archive/2020/06/scientists-predicted-coronavirus-pandemic/613003/>
- Shapiro, J. (2020, December 21). Oregon Hospitals Didn't Have Shortages. So Why Were Disabled People Denied Care? National Public Radio. Retrieved August 8, 2021, from <https://www.npr.org/2020/12/21/946292119/oregon-hospitals-didnt-have-shortages-so-why-were-disabled-people-denied-care>
- Spector-Bagdady, K., Laventhal, N., Applewhite, M., Firn, J. I., Hogikyan, N. D., Jaggi, R., Marks, A., McLeod-Sordjan, R., Parker, L. S., Smith, L. B., Vercler, C. J., & Shuman, A. G. (2020). Flattening the Rationing Curve: The Need for Explicit Guidelines for Implicit Rationing during the COVID-19 Pandemic. *The American Journal of Bioethics : AJOB*, 20(7), 77–80. <https://doi.org/10.1080/15265161.2020.1779409>
- Stone J. R. (2020). Social Justice, Triage, and COVID-19: Ignore Life-years Saved. *Medical care*, 58(7), 579–581. <https://doi.org/10.1097/MLR.0000000000001355>.
- Supady, A., Curtis, J. R., Abrams, D., Lorusso, R., Bein, T., Boldt, J., Brown, C. E., Duerschmied, D., Metaxa, V., & Brodie, D. (2021). Allocating scarce intensive care resources during the COVID-19 pandemic: practical challenges to theoretical frameworks [PDF]. *The Lancet*, 9(4), 430-434.
- University of California Critical Care Bioethics Working Group. (2020, June 17). *Allocation of Scarce Critical Resources under Crisis Standards of Care*. University of California.
- Vinay, R., Baumann, H., & Biller-Adorno, N. (2021). Ethics of ICU triage during COVID-19 [PDF]. *British Medical Bulletin*, 138(1), 5-15.



Myanmar Risk Analysis: McDonald's Case Study

Catherine He

Author Background: *Catherine He grew up in China and Canada and currently attends Ridley College in St. Catharines, Ontario in Canada. Her Pioneer research concentration was in the field of international relations and titled "Global Risk and Prediction."*

Client

McDonald's, an American fast-food chain with various international locations.

Client Proposal

While McDonald's has established a large presence in Asia, the restaurant franchise has yet to expand into Myanmar. This nation that has experienced one of the fastest growing economic expansions on the continent with service sectors gaining increasing importance and a growingly affluent population (Emerhub). Given KFC's 2015 successful expansion into the country as the first American fast-food chain to do so, McDonald's is contemplating expanding their operations into Myanmar (Nitta). However, the client is concerned that recent events, most notably the formation of a stratocracy alongside the COVID-19 pandemic, have made expansion too risky.

Introduction

Myanmar, a nation of over 54 million people, has had a turbulent year. Between the devastating COVID-19 pandemic, a parliamentary election with disputed legitimacy, and a military coup, the Southeast Asian nation is currently highly unstable. The country has a history of military rule and one of the weakest GNI per capita via the Atlas Method in the region (World Population Review). Therefore, there are several risks worth considering for restaurant expansion.

The key risks faced by Myanmar span across four main areas: political, economic, social, and geopolitical. Politically, Myanmar faces corruption without meaningful counter-initiatives, mass daily protests against the new regime, and declining democracy tightening regulations. Economically, the nation is fighting two crises at once that has exacerbated existing vulnerabilities, such as reliance on agriculture, putting the country on the brink of collapse. Socially, Myanmar faces heightened religious and ethnic tensions, with the Rohingya crisis remaining a major point of instability. Geopolitically, it faces tensions with

neighboring Bangladesh over Rohingya refugees and is the proxy-war site between China-US and China-India due to its situation between two of the largest Asian economies. Through analysis, it can be determined that political risk is the greatest threat to Myanmar's stability, which is largely due to recent political developments that have overhauled the governance of the nation and resulted in mass social discontent and violence.

Over the past decade, Myanmar was finally taking strides towards democracy. The first civilian rule of Aung San Suu Kyi at the end of 2011 helped revitalize the economy and strengthen diplomatic relations. However, the country has faltered in recent years. The 2017 Rohingya crisis prompted international condemnation, and on February 1, 2021, after approximately one decade of civilian rule, the Tatmadaw – Myanmar's ruling military junta – ended its experiment in democracy and captured full control. Rampant corruption and a looming military presence in politics have continued to be key political risks that have plagued the nation. These factors generate great uncertainty surrounding the future of business, specifically with regards to foreign expansion. Therefore, while Myanmar faces several varied key risks, developments like the military coup mean that the next major flashpoint will likely be political.

Time Horizon

This analysis is based on conditions as they existed in August 2021. The time horizon for this risk report is five months, with the analysis being highly relevant up until February 2022, one year since the beginning of the coup.

Political Risk: High

Failure of Democratic Institutions – High

On February 1, 2021, the Tatmadaw, the official name of the Myanmar armed forces, seized control via a coup d'état. Under military commander-in-chief Min Aung Hlaing's ruling, many have been detained, including Suu Kyi and National League for Democracy (NLD) members (Cuddy).

The key risk is the loss of progress towards democracy and freedom. There is a steady decline in Myanmar's Freedom House reports – from 'partly free' at 32/100 to 'not free' at 28/100 in the span of five years, with expectations of further decline. Myanmar faces disputed election results and surveilled discussion, and regulations have tightened during the coup as the military attempts to preserve its image by prohibiting internet posting about the protests (Freedom House). Free press has been overtaken by state-controlled media, a prime example being the *Myanmar Times* news outlet's silence since February 2021.

The coup has little credibility amongst the people, who continue fighting for democracy. Under promised military rule until 2023, much of the progress made by the NLD will be lost, and it is highly likely that civilian liberties will continue to decline as the crisis grows increasingly violent. For McDonald's, there

remains the possibility of nationalization of the private sector similar to what transpired after the coup of 1962, in which case all capital investments would be forfeited (McCarthy 10). The Tatmadaw has shown previously its desire to eliminate alien influence and promote state ownership, which could be incredibly detrimental to the success of McDonald's, being a foreign franchise (Cook 117). Democratic institutions in Myanmar have failed.

Civilian Protestors – High

Myanmar's civilians have been pushing against the regime in daily peaceful protests named the largest since the Saffron Revolution, presenting a massive risk to political stability (Marlow et al.). Demonstrations, generally youth-led, have prevailed despite violence (Lone et al.). The protests have drawn a wide array of people including civil servants, healthcare professionals, transport workers, and artists, predominantly in city areas such as Yangon, the largest city in Myanmar (Solomon and Mandhana). As of August 26, at least 1,031 people have been killed by the Tatmadaw since the coup began (Lone et al.).

Unrest is a high risk due to increased tensions between state and society. Myanmar's junta is responding to protestors with violence effectively amounting to war crimes, and it will likely escalate in the next few months as the opposing agendas continue clashing (Human Rights Watch). These circumstances are fueling a sharp decline in citizen mental health, as well as a credibility and productivity decrease from a paralyzed administration (Mendelson). Stability becomes unattainable as an unhappy population continues demonstrating. Given that social unrest often results in sharp contractions in the sectoral and demand dimensions as well as capital damage, McDonald's would be vulnerable and exposed to significant financial risk by expanding into the likely first destination of a major city such as Yangon (International Monetary Fund).

Political Corruption – Medium

Within Myanmar, corruption runs rampant. There have been government initiatives aimed at curbing this but without meaningful results. In military-ruled Myanmar, the Anti-Corruption Commission found that Suu Kyi accepted bribes and misused authority to gain advantageous terms in deals (Peck). Although case legitimacy is unknown, it proves the risk of political corruption in a fragile nation.

The privatization of state-owned companies and other economic reforms recently have allegedly benefited relations of senior officials (Freedom House). One example is military conglomerates, a significant source of revenue for the Tatmadaw, disproportionately benefiting higher ranking Tatmadaw officers and institutions through exploiting broad policies to engage in illicit activities for personal gain (McCarthy 18). There is too much impunity and little accountability, as the junta is more concerned about revenue and power than reputation (Reed).

The government does not operate with transparency, which is a risk to political stability because institutional capacity is utterly eroded if procedures and regulations are disregarded. Undermined legitimacy creates a lack of trust

between society and the state, translating to more uprisings as the situation worsens in the timeframe. For McDonald's, this could mean corrupt and vague policies that create uncertainty, making it difficult to conduct regular operations. The factors make Myanmar a less attractive business environment.

Shock Event: Civil War

Myanmar is incredibly politically polarized, which has created the perfect breeding ground for a civil war against the Tatmadaw. The ethnic armed organizations (EAOs), who are against the Tatmadaw's violence towards civilians, have threatened ceasefires as well as offered training to civilians for personal protection. Meanwhile, a group of deposed politicians have formed a provisional government and have approached insurgents about forming a "federal army" (The Economist). The April 16, 2021, establishment of the "National Unity Government," a government in exile formed by ousted politicians, is calling on civilians to revolt against the Tatmadaw goes to further illustrate how a civil war could occur (@NUGMyanmar, *Statement of the #NationalUnityGovernment*). After months of protesting to little effect, protestors are beginning to take up arms for democracy (Fishbein et al.). With an already fragile relationship between the state and society, all it will take is one incident for the situation to escalate into a civil war. The damage to infrastructure, decline in human capital, disruption to normal economic activity, rise in debt, and societal tension would be a devastating situation for McDonald's business operations.

Economic Risk: High

The Double Crises' Effect – High

COVID-19 has wreaked havoc in Myanmar, one of the least developed nations regionally. Myanmar had a GDP of 76.19 billion USD in 2020 and is projected to contract around 18% in Myanmar's 2021 Fiscal Year. Many industries have been disrupted, including tourism, which made up 6.7% of the GDP pre-pandemic (Knoema). With the fact that a business's revenue growth is typically influenced by the level of GDP growth within a country, contraction is the worst-case scenario for entering a foreign market.

The coup has intensified the effects of COVID-19; the country's economy is approximately 30% smaller than it would have been in the absence of the combined crises (World Bank, *Myanmar Economy*). Supply chains and shipping lines have unraveled between closures and shortages (Whelan). The US and EU have imposed targeted sanctions on major military conglomerates, causing foreign investment to stall due to the instability and staff shortages, employee safety concerns, and reputational risk (AM Best). Shortages would render a restaurant business like McDonald's, which is heavily reliant on importing fresh goods, practically incapable of standard operation and pose a threat to the company's sustainability in Myanmar. There will also be continued economic pressure through inflation and inaccessibility as the military regime hinders

growth, all of which would pose a threat to the company's sustainability in Myanmar. Foreign investors will continue pulling out, and Myanmar's economic diversity will dwindle, which is a bad sign for McDonald's expansion. While the Tatmadaw claims to be revitalizing the economy, they likely won't have accomplished much by the end of the first year of the coup.

An Insecure and Vulnerable Economy – High

Half of the population is employed in the agriculture industry that accounts for over one-fifth of the GDP (AM Best). This dependence is a high risk because Myanmar is in a hazard-prone zone, making the country susceptible to price volatility due to negative demand and supply side shocks (Steinberg et al.). The impact of a disaster on the economy would be devastating. Given that Myanmar's monsoon season is underway, there is a possibility of a disaster occurring soon that could worsen the dire economic situation.

Furthermore, the gender and ethnic exclusion, widespread underdevelopment, and weak service delivery means that human capital is insecure in Myanmar, and recovery will be slow (World Bank, *Myanmar Overview*). In fact, women's labor participation rate has been steadily declining (Statista Research Department). The banking system is on the brink of crisis (Rieffel), and foreign investors have been reassessing their involvement in the country, increasing pressure that will affect growth and stability ("Almost Half of Myanmar").

Myanmar is also the second largest producer of opium in the world and is at risk of being first place in the illegal drug trade in coming months. The harm created is incalculable, and this illegal industry creates an additional source of insecurity within the fragile economy (Reed and Jirenuwat). Economic vulnerability and insecurity poses a threat to McDonald's prospects of establishing a long-term stake in the nation.

Poverty – High

Myanmar ranks 61st on the Multidimensional Poverty Index (UNDP, *Multidimensional Poverty*), and nearly one-third of children in Myanmar live in poor households (UNICEF Myanmar). The combined coup-COVID effect could push up to 12 million people into poverty, resulting in as many as 25 million people living below the national poverty line by early 2022. Myanmar is approaching economic collapse (UNDP, *COVID-19*).

Food insecurity is rising sharply, undermining the ability of poor families to feed themselves. The average price of rice has risen 5%, with increases as high as 11% in Kachin (World Food Programme). Civilians have become desperate for money while businesses are struggling to supply themselves, which would be especially problematic for restaurant services such as the client (Bo).

The rise in poverty presents a massive risk. National security will decline with rising crime rates as more of the population becomes jobless. The increase in impoverishment means slow growth rates and the double crises effect will be prolonged, straining a new foreign business that hinges on preliminary market success. The deteriorating living spaces resulting from this could become the breeding ground for another health catastrophe. Half of the population could be

living in poverty by February of 2022, which would be devastating in all aspects. Mass poverty reduces productivity in terms of the available human capital, as well as prevents the robust rates of consumption from the middle class that businesses need to thrive, especially when entering a foreign market where civilians are unfamiliar with the company.

Social Risk: Medium

Ethnic Minority Violence – High

Kayin State, Myanmar, is home to the world's longest running civil war of the Karen minority nationalists (Rieffel). There are a variety of ethnic armed organizations representing different groups, with one example being the Arakan Army in Rakhine State. Most rebels are pushing for more autonomy ("Myanmar: Timeline"). The Tatmadaw has seemingly used COVID-19 to ramp up military pressure on ethnic minorities (Gerin), and the coup has created a resurgence of armed conflicts in the borderlands with ethnic minorities such as the Karen, the Shan, and more (Steinberg et al.). Ethnic militias have threatened ceasefires in reaction to the junta's actions (Wongcha-um).

Violence could escalate as resentment grows from more lack of autonomy under military reign, translating to security threats. Ethnic rebel groups are beginning to offer training to civilian protestors and even forming connections with the National Unity Government, who are famously anti-coup, meaning that EAO violence could escalate in the next few months as more civilians take up arms (@NUGMyanmar, #nug collectively).

The Rohingya Muslim Genocide – Medium

The Rohingya Muslims are a religious minority in Myanmar known to be the most persecuted group in the world (Faye 1). They are subject to discrimination through "Muslim-free" villages and carrying "white cards," identification for stateless people. Without citizenship, exclusion is the norm.

In 2017, Myanmar's military campaign forced more than 700,000 Rohingya to flee the nation. Today, over one million Rohingya refugees are stranded in Bangladesh, and approximately 130,000 Rohingya Muslims have been placed in squalid camps in Rakhine State (Freedom House). The UN has officially named the Rohingya Muslim crisis a genocide (Human Rights Council), an expanding into a genocide would be an undesirable situation when considering the reputation of McDonald's – a company that gives back to the community.

This crisis is a medium risk due to a socially unstable landscape of anger and despair (Diamond). Furthermore, in response to previous attacks on the Rohingya, there have been vengeful tactics used by militant groups in response, including bomb attacks (Bashar 16). However, the Rohingya have been granted access to COVID-19 vaccines, which could be a sign of good progression ("Myanmar Says Rohingya Minority"). Still, Rohingya discrimination is military in nature and will likely increase slightly, though probably not in severity, in

coming months as the Tatmadaw focuses on gaining control of the nation and grappling with protestors.

Religious Extremism – Medium

Buddhist values are widespread in a nation where nine-tenths of the population follows Theravada Buddhism (Steinberg et al.), and monks have enormous societal sway (Robertson). Despite this demographic majority, there is a fear that Buddhism is in danger of being erased, resulting in extreme defense of religion (Gunasingham 1). The military junta is backed by many radical monks and nationalists and is looking to take advantage of the deep-rooted Buddhist nationalism in Myanmar (Hardig and Sajjad).

Islamist extremism is also emerging in Myanmar. In recent years, there have been attacks by Rohingya fighters resulting from dissatisfaction with continued discrimination (Muang). Between squalid camps and displacement, Islamist extremism could escalate amid Myanmar's already-violent scenery.

The current instability offers an opening for religious minorities to be radicalized, which would further divide society when unrest and violence is already soaring. These circumstances present an unstable societal landscape that is detrimental in providing an element of unpredictability with regards to business in the country. With the resurgence in Buddhist nationalism and having a ruling Buddhist value-aligned military aiming to preserve its central role in politics, there will likely be an uptick in extremism in coming months.

Geopolitical Risk: Medium

Bangladesh – Medium

The Rohingya crisis has dealt an irrefutable blow to Myanmar's diplomatic relationship with Bangladesh. Since 2017, the neighboring nation has hosted nearly 1.1 million Rohingya that pose increasing challenges for the country (Sharma, *Indefinite Hosting*). Bangladeshi officials are doubting Myanmar's willingness to take back the displaced peoples that they do not officially recognize as refugees and have been appealing for worldwide pressure on Myanmar (Banerjee 4). However, the Rohingya are not the only reason for the risk of growing tensions between the two countries. The porous border between Myanmar and Bangladesh provides suitable terrain for insurgent operations by radical groups, which could lead to conflict as uprisings increase due to the junta (Bashar 17).

The military coup has undoubtedly increased fear among the Rohingya refugees about returning to Myanmar due to the history of military discrimination. Therefore, it should be expected that the countries' relations will continue to deteriorate, affecting trade and diplomacy. It is unlikely that much will progress within the next few months, but the stateless refugees serve as a major irritant that provide instability. Should there be an event that increases the number of displaced people similar to the incident in 2017, it may be the final straw.

US-China Proxy War – Medium

China, Myanmar's extremely powerful neighbor, has been a long-time supporter of the nation and plays a dominant role in the economy through the Belt and Road Initiative (BRI). This initiative aims to increase China's foreign connectivity through investing in infrastructure and opening transport and economic corridors (Mobley 52). The BRI is an effort to gain dominance in the Asia-Pacific to serve commercial and military purposes. The rift between Myanmar and the international community due to the coup is allowing China to build closer ties with the nation to prime for economic exploitation (Gong 2). Overreliance on a single nation is risky enough on its own, especially if the relationship sours. In Myanmar's case, it is already experiencing strain to Sino-Burmese relations under the new regime – the large-scale protest of Myanmar civil society against the Myitsone dam project due to adverse environmental impacts forced the Myanmar military government to suspend the project (Gong and Balazs 110).

President Biden has made it very clear that his foreign policy consists of promoting democracy while vetoing China to ensure US hegemony. Therefore, the US is in a position where any tougher countermeasures would drive Myanmar closer to China, whereas no action will make it obvious that they don't have hegemony (Yu). After Biden's controversy regarding Afghanistan, it is possible that he will act in Myanmar as redemption, which could be a crackdown from the US through increased sanctions or even military presence in the next few months, preventing national growth and development and hindering McDonald's business operations.

India-China Proxy War – Low

As China and India mount troops along Myanmar's contested northern border, the long-time rivals have been brought close to conflict. Myanmar, firmly stuck in between the two Asian giants, could soon be pulled in countervailing directions as a power tussle over territory unfolds. India is doing everything it can to hinder China's growth, including silence regarding the military coup so that China doesn't become Myanmar's last resort (Sharma, *India's Silence*). China appeared when Myanmar became a pariah to promote the Belt and Road Initiative, which obviously is opposite of India's desire as the two giants compete to build infrastructure along the Line of Actual Control (Loomba 100). As the fight for dominance grows heated in a recovering world, Myanmar may be at risk of border violence in the near future, which would cause immobilization and widespread violence with the military junta already trying grasp control over the civil disobedience movement and ethnic rebels. This risk of incapacitated borders poses a huge threat to the future of McDonald's in Myanmar. Being an American company that relies on imports, business would likely crumble under these circumstances.

Recommendation

Expansion into unfamiliar territory is daunting on its own, let alone expanding into a nation of instability. McDonald's could potentially damage its reputation by prematurely developing a Myanmar location during the COVID-19 pandemic and a military coup where thousands are dying. As violence roams the streets in major cities that are favorable locations for expansion, McDonald's could face physical damage and threats to employee safety. These are adverse for such a large international franchise known for giving back to the community. Under these circumstances, there is a significant risk of supply chain disruptions that could limit access to necessary supplies and render the restaurant chain unable to operate. Given that businesses are constantly closing operations in strikes against the military regime, it seems as though there is no substantial financial benefit of expansion (Al Jazeera, *Protestors Rally*). Other countries, predominantly the United States, have imposed numerous sanctions on Myanmar which could be troublesome given that McDonald's is an American restaurant chain. Sanctions also produce tremendous financial pressure by raising production costs and a less than ideal operating environment. Furthermore, businesses are at risk of nationalization as the junta attempts to control assets amid a tumultuous economic crisis and criticism from the international community – this could even occur as the government endeavors to assert its power over foreign industries. McDonald's could potentially lose all gains should this be the case, negatively impacting the brand's reputation. Rampant corruption and a lack of government transparency could result in legal missteps. At the same time, societal divisions cause fear and uncertainty that is known to hamper businesses as they consider expanding operations into uncharted territory.

Therefore, it is highly recommended that McDonald's delay and potentially reconsider their expansion into Myanmar for now. A time that McDonald's could re-evaluate expansion would be in 2023 after the coup is projected to be over and elections begin, allowing for a more stable political environment; however, this is subject to the changing desires of the junta and whether they follow through on their plan. Currently, Myanmar is facing impending threats of civil war, daily civilian protests endangering the legitimacy of the junta, mass poverty that limits human capital and consumer expenditure, and a crumbling economy that continues to be reliant on the vulnerable agriculture industry, among other issues. As a result, expanding McDonald's into Myanmar now would be too risky for the brand's image and unsustainable for revenue and profits. The barriers to expansion are too high at this point and McDonald's risks losing more than it stands to gain.

Works Cited

Albert, Eleanor, and Lindsay Maizland. "What Forces Are Fueling Myanmar's Rohingya Crisis?" *Council on Foreign Relations*, 23 Jan. 2020, www.cfr.org/backgrounder/rohingya-crisis. Accessed 20 Aug. 2021.

- Al Jazeera. "Protesters Rally, Businesses Close in Strike against Myanmar Coup." *Al Jazeera*, 22 Feb. 2021, www.aljazeera.com/news/2021/2/22/myanmar-new. Accessed 19 Jan. 2022.
- AM Best. *Best's Country Risk Report Myanmar*. AM Best, 2021, <https://www3.ambest.com/ratings/cr/reports/myanmar.pdf>. Accessed 28 Aug. 2021.
- "Almost Half of Myanmar Risks Falling into Poverty by 2022: UNDP." *Al Jazeera* 30 Apr. 2021, www.aljazeera.com/economy/2021/4/30/almost-half-of-myanmar-risks-falling-into-poverty-by-2022-undp. Accessed 10 Aug. 2021.
- Banerjee, Sreeparna. "The Rohingya Crisis and its Impact on Bangladesh-Myanmar Relations." *ORF Issue Brief No. 396*, Aug. 2020, https://www.orfonline.org/wp-content/uploads/2020/08/ORF_IssueBrief_396_Bangladesh-Myanmar.pdf.
- Bashar, Iftekharul. "Rohingya Crisis and Western Myanmar's Evolving Threat Landscape." *Counter Terrorist Trends and Analyses*, vol. 11, no. 6, International Centre for Political Violence and Terrorism Research, 2019, pp. 14–18, <https://www.jstor.org/stable/26662256>. Accessed 19 Jan. 2022.
- Bo, Muang. "Covid and Military Coup Cripple Myanmar Economy." *DW*, 26 Aug. Aug. 2021, www.dw.com/en/covid-and-military-coup-cripple-myanmar-economy/a-58992906. Accessed 29 Aug. 2021.
- Cook, Paul. "Policy Reform, Privatization, and Private Sector Development in Myanmar." *South East Asia Research*, vol. 2, no. 2, IP Publishing Ltd, 1994, pp. 117–40, <http://www.jstor.org/stable/23746861>. Accessed 15 Jan. 2022.
- Cuddy, Alice. "Myanmar Coup: What Is Happening and Why?" *BBC News*, 1 Apr. 2021, <https://www.bbc.com/news/world-asia-55902070>. Accessed 1 Aug. 2021.
- Diamond, Cape. "Myanmar Rohingya Feel Anger and Despair before Election." *Nikkei Asia*, 30 Oct. 2020, <https://asia.nikkei.com/Politics/Myanmar-Rohingya-feel-anger-and-despair-before-election>. Accessed 29 Aug. 2021.
- Emerhub. "Top 8 Reasons to Invest in Myanmar." *Emerhub*, 3 Jan. 2019, <https://emerhub.com/myanmar/top-reasons-to-invest-in-myanmar/>. Accessed 13 Jan. 2021.
- Faye, Malang. "A forced migration from Myanmar to Bangladesh and beyond: humanitarian response to Rohingya refugee crisis." *Journal of International Humanitarian Action*, vol. 6, no. 13, 2021, <https://doi.org/10.1186/s41018-021-00098-4>. Accessed 13 Jan. 2022.
- Fishbein, Emily, et al. "'Our Only Option': Myanmar Civilians Take up Arms for Democracy." *Al Jazeera*, 15 Jul. 2021, <https://www.aljazeera.com/news/2021/6/15/our-only-option-myanmar-civilians-take-up-arms-for-democracy>. Accessed 25 Aug. 2021.
- Freedom House. "Myanmar: Freedom in the World 2021 Country Report." *Freedom House*, 3 Mar. 2021, <https://freedomhouse.org/country/myanmar/freedom-world/2021>. Accessed 15 Aug. 2021.

- Gerin, Roseanne. "Arakan Army Accuses Myanmar of Using COVID-19 to Ramp up Military Pressure." *Radio Free Asia*, 11 Oct. 2020. <https://www.rfa.org/english/news/myanmar/military-pressure-09032020214958.html>. Accessed 1 Sep. 2021.
- Gong, Xue, and Daniel Balazs. "Emerging Soft Turn in China's Energy Security Cooperation with Southeast Asia." *China Review*, vol. 21, no. 4, The Chinese University of Hong Kong Press, 2021, pp. 109–40, <https://www.jstor.org/stable/48635894>. Accessed 17 Jan. 2022.
- Gong, Xue. "Understanding the Belt and Road Initiative in Myanmar: A Socio-Politico and Economic Approach." *China and the World*, vol. 3, no. 04, 2020, <https://doi.org/10.1142/s2591729320500169>. Accessed 2 Sep. 2021.
- Gunasingham, Amresh. "Buddhist Extremism in Sri Lanka and Myanmar: An Examination." *Counter Terrorist Trends and Analyses*, vol. 11, no. 3, International Centre for Political Violence and Terrorism Research, 2019, pp. 1–6, <https://www.jstor.org/stable/26617827>. Accessed 28 Aug. 2021.
- Hardig, Anders C, and Tazreena Sajjad. "The Military Coup in Myanmar Presents Opportunities to Buddhist Nationalists." *The Conversation*, 23 Aug. 2021, <https://theconversation.com/the-military-coup-in-myanmar-presents-opportunities-to-buddhist-nationalists-154459>. Accessed 23 Aug. 2021.
- Human Rights Council. *Report of the Independent International Fact-Finding Mission on Myanmar*. Human Rights Council, 2018, https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/39/64.
- Human Rights Watch. "Myanmar: Coup Leads to Crimes against Humanity." *Human Rights Watch*, 3 Aug. 2021, <https://www.hrw.org/news/2021/07/31/myanmar-coup-leads-crimes-against-humanity>. Accessed 17 Aug. 2021.
- International Monetary Fund. "Could Renewed Social Unrest Hinder the Recovery?" *IMF Blog*, 1 Nov. 2021, blogs.imf.org/2021/07/13/could-renewed-social-unrest-hinder-the-recovery/. Accessed 17 Jan. 2022.
- Knoema. "Myanmar Contribution of Travel and Tourism to GDP (% of GDP), 1995-2019." *Knoema*, knoema.com/atlas/Myanmar/topics/Tourism/Travel-and-Tourism-Total-Contribution-to-GDP/Contribution-of-travel-and-tourism-to-GDP-percent-of-GDP. Accessed 17 Jan. 2022.
- Loomba, Deepak. "Coagulating The Line Of Actual Control Between India And China." *World Affairs: The Journal of International Issues*, vol. 24, no. 4, Kapur Surya Foundation, 2020, pp. 98–111, <https://www.jstor.org/stable/48609812>. Accessed 19 Jan. 2022.
- Lone, Wa, et al. "Protest Drove a Generation to Myanmar's Streets. Military Crackdowns Sent Hundreds to Their Deaths." *Reuters*, August 27, 2021. <https://graphics.reuters.com/MYANMAR-POLITICS/zjppqkkoaqpx/>. Accessed 28 Aug. 2021.
- Marlow, Iain, et al. "Myanmar Sees Biggest Protest in Years as Coup Opposition Grows." *Bloomberg*, 7 Feb. 2021. <https://www.bloomberg.com/news/articles/2021-02-07/myanmar-protesters-defy-generals-to-march-against-military-coup>. Accessed 17 Aug. 2021.
- McCarthy, Gerard. *Military Capitalism in Myanmar: Examining the Origins, Continuities and Evolution of Khaki Capital*. Singapore, ISEAS Yusof Ishak Institute, 2019.

- Mendelson, Allegra. "Coup, COVID Take Toll on Young People's Mental Health in Myanmar." *Al Jazeera*, 25 Aug. 2021, <https://www.aljazeera.com/news/2021/8/25/coup-covid-takes-toll-on-young-peoples-mental-health-in-myanmar>. Accessed 26 Aug. 2021.
- Mobley, Terry. "The Belt and Road Initiative: Insights from China's Backyard." *Strategic Studies Quarterly*, vol. 13, no. 3, Air University Press, 2019, pp. 52–72, <https://www.jstor.org/stable/26760128>. Accessed 17 Jan. 2022.
- Muang, Manny. "Myanmar: A New Muslim Insurgency in Rakhine State." *Crisis Group*, 18 Jan. 2018, <https://www.crisisgroup.org/asia/south-east-asia/myanmar/283-myanmar-new-muslim-insurgency-rakhine-state>. Accessed 19 Jan. 2022.
- "Myanmar Says Rohingya Minority Will Get Covid-19 Vaccine." *VOA*, 1 Sep. 2021, https://www.voanews.com/a/covid-19-pandemic_myanmar-says-rohingya-minority-will-get-covid-19-vaccine/6210048.html. Accessed 2 Sep. 2021.
- "Myanmar: Timeline of a Fragile Democracy." *Al Jazeera*, 1 Feb. 2021, <https://www.aljazeera.com/news/2021/2/1/myanmar-military-rule-to-fragile-democracy>. Accessed 17 Aug. 2021.
- Nitta, Yuichi. "KFC seeks second serving in Myanmar with Mandalay expansion." *Nikkei Asia*, 24 May 2017, <https://asia.nikkei.com/Business/KFC-seeks-second-serving-in-Myanmar-with-Mandalay-expansion>. Accessed 12 Sep. 2021.
- @NUGMyanmar. "Statement of the #NationalUnityGovernment of #Myanmar on the 4th anniversary of the atrocity crimes committed against the #Rohingya people in 2017." *Twitter*, 24 Aug. 2021, 6:34 p.m., <https://twitter.com/NUGMyanmar/status/1430297406923448324>. Accessed 24 Aug. 2021.
- @NUGMyanmar. "#nug collectively with the Ethnic Armed Organisations (#EAO) has formed a committee of implementing the Peoples Embrace Project to welcome and embrace those #Soldiers and #Police who have defected to take part in the #CivilDisobedienceMovement." *Twitter*, 19 Sep. 2021, 3:30 a.m., <https://twitter.com/NUGMyanmar/status/1439492078564569090>. Accessed 19 Sep. 2021.
- Peck, Grant. "Myanmar Anti-Corruption Body Files Cases Against Suu Kyi." *Associated Press*, 10 Jun. 2021. <https://apnews.com/article/myanmar-health-coronavirus-pandemic-government-and-politics-4bca7a9af54ad1bb870feac08f5ef023>. Accessed 24 Aug. 2021.
- "Population, total - Myanmar." *The World Bank*, 2020, <https://data.worldbank.org/indicator/SP.POP.TOTL?locations=MM>. Accessed 12 Sep. 2021.
- Reed, John and Ryn Jirenuwat. "How Myanmar Coup Fuelled Rise in Illegal Drugs Trade." *Financial Times*, 22 Aug. 2021, <https://www.ft.com/content/e9ee221-cba4-483d-8799-c2a356af4cb4>. Accessed 4 Sep. 2021.
- Rieffel, Lex. "Peace and War in Myanmar." *Brookings*, 6 Dec. 2019, <https://www.brookings.edu/blog/order-from-chaos/2019/12/06/peace-and-war-in-myanmar/>. Accessed 2 Sep. 2021.
- Robertson, Holly. "Buddhist Extremism: Meet the Violent Followers of a Religion Widely Known for its pacifism." *ABC News*, 20 Oct. 2018, <https://www.abc.net.au/news/2018-10-21/buddhist-extremism-meet-the-religions-violent-followers/10360288>. Accessed 14 Sep. 2021.

- Sharma, Indrajit. "Indefinite Hosting of Rohingya Refugees a Growing Concern for Bangladesh." *The Diplomat*, 10 Jul. 2021, <https://thediplomat.com/2021/07/indefinite-hosting-of-rohingya-refugees-a-growing-concern-for-bangladesh/>. Accessed 25 August 2021.
- Sharma, Kiran. "India's Silence toward Myanmar Shows Its Wariness of China." *Nikkei Asia*, 25 Feb. 2021, <https://asia.nikkei.com/Spotlight/Myanmar-Crisis/India-s-silence-toward-Myanmar-shows-its-wariness-of-China>. Accessed 7 Aug. 2021.
- Solomon, Feliz, and Niharika Mandhana. "Myanmar Coup: What's Happening and Why Are There Protests?" *The Wall Street Journal*, 15 Mar. 2021, <https://www.wsj.com/articles/myanmar-coup-aung-san-suu-kyi-11612169495>. Accessed 4 Aug. 2021.
- Statista Research Department. "Myanmar: Labor Force Participation Rate by Gender 2018." *Statista*, 9 Nov. 2021, www.statista.com/statistics/1066993/myanmar-labor-force-participation-rate-by-gender/. Accessed 18 Jan. 2022.
- Steinberg, David I. et al. "Myanmar." *Encyclopedia Britannica*, 6 Aug. 2021, <https://www.britannica.com/place/Myanmar>. Accessed 10 Aug. 2021.
- UNDP. *COVID-19, Coup D'Etat and Poverty: Compounding Negative Shocks and their Impact on Human Development in Myanmar*. UNDP, 2021, https://www.asia-pacific.undp.org/content/rbap/en/home/library/democratic_governance/covid-19-coup-d-etat-and-poverty-impact-on-myanmar.html.
- UNDP. *Multidimensional Poverty Index: Developing Countries*. UNDP, 2020, <http://hdr.undp.org/en/composite/MPI>.
- UNICEF Myanmar. *Overcoming Child Poverty in Myanmar: Investing in Human Capital, Responding to COVID-19*. UNICEF Myanmar, 2020, <https://www.unicef.org/myanmar/media/5806/file/Overcoming%20Child%20Poverty%20in%20Myanmar.pdf>.
- Whelan, Sam. "Supply Chains Unravel as Myanmar Military Escalates Retaliation to Protests." *The Loadstar*, 11 Mar. 2021, <https://theloadstar.com/supply-chains-unravel-as-myanmar-military-escalates-retaliation-to-protests/>. Accessed 15 Aug. 2021.
- The Economist. "Who Are Myanmar's Ethnic Militias?" *The Economist*, 15 Apr. 2021, <https://www.economist.com/the-economist-explains/2021/04/15/who-are-myanmars-ethnic-militias>. Accessed 28 Aug. 2021.
- Wongcha-um, Panu. "Myanmar Coup Threatens Ceasefire, Ethnic Rebel Leader Says." *Reuters*, 3 Feb. 2021, <https://www.reuters.com/article/us-myanmar-politics-insurgency-idUSKBN2A31JO>. Accessed 25 Aug. 2021.
- World Bank. "Myanmar Economy Expected to Contract by 18 Percent in FY 2021: Report." *World Bank*, 26 Jul. 2021, <https://www.worldbank.org/en/news/press-release/2021/07/23/myanmar-economy-expected-to-contract-by-18-percent-in-fy2021-report>. Accessed 20 Aug. 2021.
- World Bank. "Myanmar Overview." *World Bank*, 2020, <https://www.worldbank.org/en/country/myanmar/overview>. Accessed 27 Aug. 2021.

- World Food Programme. *Myanmar Market Price Update*. World Food Programme, March 2021, <https://reliefweb.int/sites/reliefweb.int/files/resources/Market%20Price%20Update%20WFP%20Myanmar%20%28March%202021%29.pdf>.
- World Population Review. "Poorest Asian Countries 2021." *World Population Review*, worldpopulationreview.com/country-rankings/poorest-asian-countries. Accessed 17 Jan. 2022.
- Yu, Kaho. "Myanmar Coup: Prospect of Sanctions and ESG Exposure Rattles Foreign Investors." *Verisk Maplecroft*, 4 Feb. 2021, <https://www.maplecroft.com/insights/analysis/myanmar-coup-prospect-of-sanctions-and-esg-exposure-rattles-foreign-investors/>. Accessed 4 Aug. 2021.



A Contemporary Perspective on the Health of the Great Barrier Reef and the Wellbeing of the Torres Strait Islanders

Chong (Phelps) Tin

Author Background: *Phelps Tin grew up in Shanghai, China, before going to the UK. He now lives in London and currently attends Eton College in Windsor, Berkshire in the United Kingdom. His Pioneer research concentration was in the field of environmental studies and titled “Global Indigenous Health.”*

Abstract

As native peoples to the Torres Strait Islands, the Islanders have lifestyles based around their surrounding ecosystem. However, with the deterioration of the Great Barrier Reef, which plays a crucial role in the culture and wellbeing of the Islanders, their deep spiritual and physical connection with the environment is at risk. The aim of this study is to better understand the relationship between the health of the reef and the health of the Torres Strait Islander community. Thus, the present and future health of the Great Barrier Reef along with the physical and mental health status of the Torres Strait Islanders are evaluated. Despite the establishment of several reasons for the seemingly unsolvable wide health disparity between Indigenous and Non-Indigenous Australians, this study presents another potential major factor—being disconnected from culture. The significant decline of the reef’s health has caused consequences from a discontinuation of traditional practices to forced migration from their ancestral land. Such consequences are detrimental to the Islanders’ mental health and can cause severe distress to the communities, challenging the survival of the Islander culture. Thus, it is crucial to start developing short term solutions and to continue decarbonizing society and protecting the Great Barrier Reef.

1. Introduction

Coral reefs, underwater ecosystems with reef-building corals, are one of Earth’s most significant biomes. Most reefs are located between the Tropics of Cancer and Capricorn in the Pacific Ocean, the Indian Ocean, the Caribbean Sea, the Red Sea, and the Persian Gulf. Sometimes called the “rainforests of the sea”,

coral reefs are the habitat of around a quarter of all marine animals whilst covering less than one percent of the ocean floor (Reaka-Kudla, 1997; Knowlton et al., 2010). Coral reefs are mostly found in relatively shallow waters where the sea surface temperature is between twenty degrees and thirty-seven degrees Celsius. The abundance of life in these ecosystems has proved essential to coastal communities globally, with estimates suggesting that over five hundred million people's livelihoods depend on the services associated with and resources provided by coral reefs (Hoegh-Guldberg et al., 2019).

Reefs have acquired an important role in the formation of society and culture by allowing reef fisheries to flourish, protecting the coastline during tropical storms, and being hotspots of tourism. Humankind has been interacting with coral ecosystems for millennia, and thus recent changes to coral reefs are also impacting the people whose lives depend on the health of the marine and coastal environment. In this study, the contemporary relationship between the Indigenous community, specifically the Torres Strait Islanders, and the Great Barrier Reef will be examined to evaluate the societal effects that the changing reef has caused so far, and potential impacts it may have on the community in the future. In doing so, this paper encourages future research to explore potential solutions to the socio-environmental issues that entrap both humans and non-human others.



Figure 1. Map of Australia, with the location of the Great Barrier Reef Marine Park represented by the green region. (Commonwealth of Australia, 2017)

The central aim of this study is to evaluate the range of impacts the deteriorating health of the Great Barrier Reef has on the Torres Strait Islanders. The hypothesis of this paper is that the deterioration of reef health will be reflected in the society as the Islanders heavily rely on the reef materially and spiritually (Markwell, 2017). In particular, this study is concerned with any changes to the mental or physical health of the Islanders as the surrounding

environment with which they are deeply connected alters. An intriguing and relevant concept to this study is “solastalgia” (Albrecht et al., 2007), which is a term originally established by environmental philosopher Glenn Albrecht. In his words, solastalgia is “the pain or sickness caused by the loss or lack of solace and the sense of isolation connected to the present state of one’s home and territory” (Albrecht et al., 2007, p. 45). Solastalgia, interpreted environmentally, can be used to describe the distress caused by forced migration or other compromise due to the changing climate. The term links the distress of an ecosystem to the resulting human distress and is a term of increasing relevance as the impacts of climate change become more pronounced and widespread. In this context, the term embodies the recent stories of the Torres Strait Islanders, whose forthcoming displacement due to climate change is entangled with that of the coral reef. The findings of this project will be crucial for further developmental agendas regarding the Islanders, as the climate will most likely continue to change dramatically, and the resilience of the Great Barrier Reef will continue to be affected. Although this paper primarily sheds light on the Indigenous experiences associated with the changing health of their oceanic environment, it will further encourage future research to investigate the impact of changing reef ecosystems on other local communities.

1.1. The Great Barrier Reef

The Great Barrier Reef, located off the north-eastern coast of Australia, is the largest reef complex in the world. It consists of 2,900 individual reefs, 800 fringing reefs, and 900 islands, spanning over 2,300 kilometers (Brodie and Pearson, 2016). First records reveal reefs growing on the continental shelf as early as the Miocene Epoch, but the reef has undergone constant accumulation and significant change to become what it is today (Hopley, Smithers, and Parnell, 2007). The reef’s foundation consists of numerous colonies of coral polyps living on calcium carbonate exoskeletons of their predecessors. Home to more than 1,500 species of fish, 134 species of sharks and rays, over 30 marine mammals, six of the seven threatened marine turtles, 411 species of hard coral, and one third of the world’s soft corals, the reef hosts one of the most biodiverse habitats on Earth (World Wildlife Fund, 2015).

There is significant scientific interest in the area, and, inevitably, it has also become a tourist hotspot, attracting people from all over the world. In addition to other factors, this makes the Great Barrier Reef an immensely valuable economic asset. Estimates from the year 2015-2016 suggest that the reef provided over 64,000 indirect jobs and generated 6.4 billion USD for Australia (Deloitte Access Economics, 2017). There is no doubt that the reef is an irreplaceable asset that benefits both Australia and the world economically and environmentally.

However, over the past several decades, the reef has undergone a series of significant transformations. Its health has declined sharply, and some surveys show that as much as 50% of coral coverage has deteriorated and disappeared (De’ath et al., 2012). The main reasons for this substantial loss include increasingly frequent and severe coral bleaching events and tropical cyclones (Hoegh-Guldberg et al., 2018). Furthermore, there is a growing population of invasive species, such as the coral-polyp-feeding Crown of Thorns starfish

(*Acanthaster planci*), which also threaten the abundance of coral (Richards and Day, 2018), a problem that is often attributed to global climate change. In recognition of the importance of the site, the Australian government has imposed tight regulations on tourism, commercial fishing, and other damaging uses of the area such as mining and oil-drilling. Moreover, efforts to reduce water pollution including sediment and chemical run-off, as well as noise pollution, are also being pursued by the Australian government (Commonwealth of Australia, 2018).

Nonetheless, many other threats are harder to control. Global ocean acidification and rising sea surface temperatures are both ongoing threats that are challenging and altering the reef (Fabricius et al., 2020). As a result of rising greenhouse gas concentrations in the atmosphere, there is an increasing amount of radiation being trapped by the planet. Reportedly, approximately 93% of the extra heat has been absorbed by the upper layers of the ocean already, leading to increased ocean temperatures (Hoegh-Guldberg et al., 2019). In coral reefs, the higher temperatures induce coral bleaching, which weakens and kills expanses of coral. Additional carbon dioxide concentration—around 30% of that produced by human activities—has also been absorbed by the upper layers of the ocean (Hoegh-Guldberg et al., 2019). Carbon dioxide forms a weak acid when dissolved in water and subsequently reduces the pH of the water. The effects include the weakening of existing carbonate coral exoskeletons and the overall reduction of carbonate ion concentration in the region. Carbonate ions are fundamental for many biological processes such as reproduction and shell formation for a variety of marine organisms (Pendleton et al., 2019). In fact, geological studies exploring geochemical proxies for pH have shown that the current rate of ocean acidification is unprecedented within the last 65 million years, if not 300 million years (Hönisch et al., 2012). As a result of both ocean acidification and increased sea surface temperatures, there is a growing list of organisms and processes affected by the rapidly changing ocean conditions (Sala and Knowlton, 2006). Such changes and challenges affecting the Great Barrier Reef ecosystem take a significant toll on human and social survival and wellbeing.

1.2. Torres Strait Islanders

Torres Strait Islanders (approximate population: 4,500 on the Torres Strait Islands plus 28,000 in mainland Australia) are one of the two distinct Indigenous groups of Australia (Australian Bureau of Statistics, 2016). Like the Aboriginal people of mainland Australia, the Torres Strait Islanders were the first human inhabitants to occupy Australia prior to the British colonization, and their society has evolved with a close relationship to their natural environment. The Torres Strait waterway lies between Papua New Guinea to the north and Queensland of the Australian mainland to the south. In the waterway, there are more than 100 islands, of which 17 are inhabited by 18 communities (Encyclopaedia Britannica, 2020). A 2010 census found that the population of islanders makes up less than 1% of the Australian population and less than 10% of the Indigenous population in Australia. The population is of Melanesian descent, similar to the people of Papua New Guinea, and hence shares many customs and cultural traits with those of Papua New Guinea. The Islanders have

inhabited the islands for over ten thousand years, but their rights to the land were taken away following the annexation of the islands by the British in 1879. This territory was subsequently placed under the control of the state of Queensland.

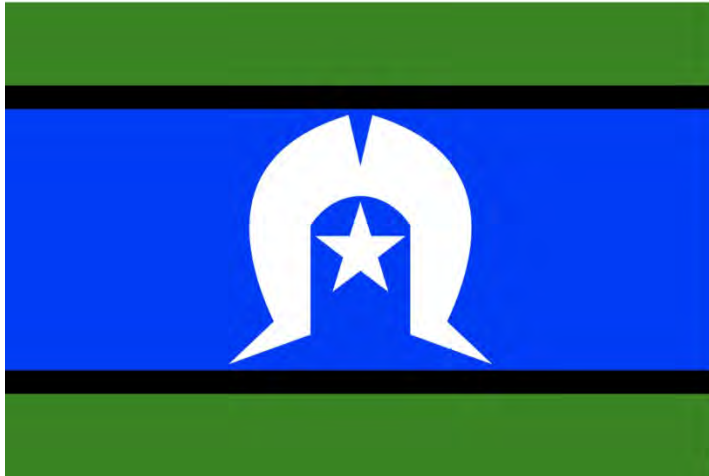


Figure 2. *Torres Strait Islander Flag (Australian Institute of Aboriginal and Torres Strait Islander Studies, n.d.)*

For thousands of years, the surrounding diverse waters and ecosystems have been the basis of resources and food, sustaining the islanders' livelihoods. Represented by the five-pointed star on the Torres Strait Islander Flag (Figure 2), the islands can be divided into five slightly different cultural groups. Moreover, the flag demonstrates the appreciation and connection the Islanders have with their natural landscape: the green is said to represent the land, just as the blue represents the sea. The black lines represent the Torres Strait Islanders, who connect the land and sea, and the white colour of the icon symbolizes peace (Australian Institute of Aboriginal and Torres Strait Islander Studies, n.d.).

Despite sharing a common way of life and reliance on the waterway, each island has geographical features that determine its supplementary means of survival. The islands that rely most on the surrounding marine environment are the central islands, also known as the fishing islands, which are abundant with diverse marine ecosystems and coral reefs. Traditionally, from surrounding ecosystems the islands relied primarily on snappers, turtles, dugongs, and sharks, but other marine organisms such as oysters were also often consumed and used (Korff, 2021). Although several islands also rely on agricultural production, mainly consisting of timber, yams, bananas, coconuts, and other fruits, most islands rely on fishing as the primary economic means. Common fisheries today include lobster, mackerel, crab, and sea cucumber fisheries; however, although traditional fishing has always included dugongs and turtles, this practice is now tightly regulated by the Torres Strait Fisheries Act (TFSA) to protect the species (Noriega, et al., 2020).

The Torres Strait Islanders have cultivated unique cultural practices of

their own, and such practices are often different from those practiced by the Aboriginal peoples of mainland Australia. Nevertheless, both believe the natural world to be a creation of their human and non-human ancestors (Mullins, 1995). Tagai, a great fisherman and spirit, is believed to be the creator of the world, and his stories shape Torres Strait Islanders' law, customs, and practices. The traditional land and nearby waterway play a fundamental role in their historical and contemporary identity, and the accumulated knowledge of the sky and sea has proven invaluable to understanding changes in the seasons, when to plant gardens and hunt for turtles or dugongs, and how to navigate the seas (Hamacher et al., 2017). As expressed by the deputy mayor of the Torres Strait Island regional council, Getano Lui, "What is instilled in us and our ancestors is [that] if the Torres Strait sinks, we'll sink with it" (Lopez, 2019). This statement highlights the urgency of improving the health of the Great Barrier Reef, as it has a strong correlation with Indigenous health and wellbeing in the times of global climate change.

The arrival of Christian missionaries impacted the Islanders significantly, resulting in the forced adoption of Christianity throughout the Torres Strait Islands in the late nineteenth century. However, the Christian missionaries are often credited the role of ending intra-island conflicts and providing protection against overexploitation of natural resources by foreign interests (Queensland Museum, n.d.).

Nevertheless, foreign influences and colonial domination often resulted in severe dispossession of land and other environmental and cultural resources of the Indigenous peoples. The occupation of the region by Christian missionaries, European settlers, and Asian fishermen had a significant impact on the culture and customs of the Indigenous peoples.

In an attempt to regulate the marine industries and prevent the further exploitation of Indigenous peoples, the Australian Government decided to annex the islands and place them under the management of the Queensland government in 1879 (Shnukal and Ramsay, 2004). The Queensland government introduced severely restrictive laws to control the islanders that essentially meant the loss of civil rights and freedom for the peoples until the 1960s. Another issue that greatly impacted the peoples is the onset of several European diseases, which eradicated a large portion of the Indigenous population on the Islands, as they did not have natural immunity to them (Korff, 2021).

It was not until 1965 that the Torres Strait Island Act of 1939 was replaced by the Aborigines and Torres Strait Island Affairs Act, which removed all the restrictions placed on the Torres Strait Islander peoples. In 1967, they would be deemed Australian citizens, and over the next few decades, there would be a persistent pursuit of self-determination and land ownership of the Torres Strait Islands (Queensland Parliament, 2015). In 2013, the Australian legislature passed the updated Aboriginal and Torres Strait Islander Peoples Recognition Act of 2013, which acknowledges Aboriginal and Torres Strait Islander peoples as the first inhabitants of Australia. Even though this was a major step towards reducing inequality, there is still a substantial disparity between Indigenous and non-Indigenous peoples in terms of healthcare services, employment, and other aspects such as food security. Such disadvantages are some of the factors that make the Indigenous peoples more vulnerable to the effects of a changing environment.

2. Literature Review

In order to accurately assess the myriad effects of the changing Great Barrier Reef on the Torres Strait Islanders, the health of both entities must be reviewed first. Then, this paper will examine other research which features the interdependence of coastal Indigenous communities and surrounding marine ecosystems. This literature review will be divided into two parts: a review concerning the health of the Great Barrier Reef, and another concerning the health of the Islanders. These two sections will be followed by my analysis and discussion that highlight the interdependence between coastal Indigenous groups and their environment.

2.1. Health of the Great Barrier Reef

One of the greatest advantages of studying the Great Barrier Reef is the extensive amount of research that has been dedicated to understanding and tracking changes of the reef over time. However, the main difficulties and limitations include some sources being out of date, and the imprecision of numbers regarding specific population sizes for even keystone species due to the large geographical expanse of the reef.

The basis of assessment of the health of the system in this study will be founded upon two primary measures: biodiversity and coral cover. It is important to evaluate the two measures separately as there is a dichotomy of evidence on whether decreased coral cover is indeed correlated to a decline in biodiversity. Whilst some studies point to an independence between habitat loss and biodiversity (Spalding and Jarvis, 2002; Doherty and Fowler, 1994), there are also many that clearly indicate the loss of coral cover directly causes the suffering of coral-specialists (organisms that are specialized to inhabit coral reefs) (Munday and Allen, 2002; Randall et al., 1990).

Most recently, a study published in 2021 by The Royal Society estimated that nearly half of reef fish richness could decline following global total coral reef loss (Giovanni et al., 2021). By assessing changes to the phylogenetic diversity and functional diversity of sites worldwide, they predicted that the long-term biodiversity of coral reefs will decline significantly with the degradation of coral reefs. However, they also acknowledged that around 25% of species had increased in population following declined coral cover. The research further stresses the importance of examining both indicators of reef ecosystem health: biodiversity and coral cover.

A study published in 2019 led by researcher Glenn De'ath, who has written extensively about the Great Barrier Reef, investigated the change in coral cover of the Great Barrier Reef over 27 years, from 1985-2012 (De'ath, 2012). As illustrated by Figure 3, the results displayed high spatial variation: the highest coral cover percentages were evident in the far northern and southern regions, whilst the lowest were in the central inshore reefs.

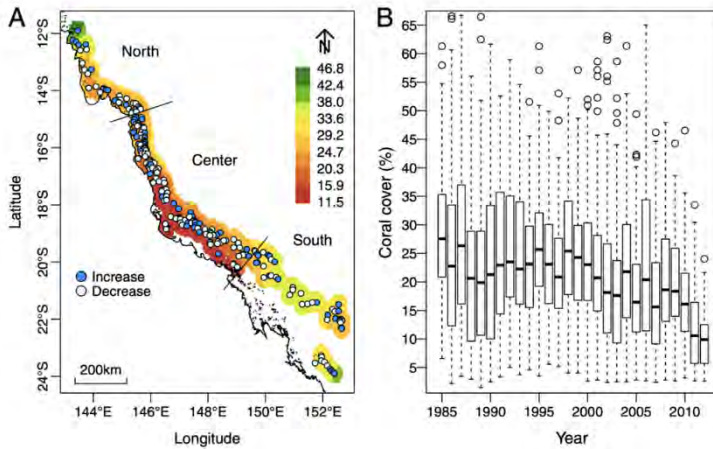


Figure 3. “Coral cover on the GBR. (A) Map of the GBR with color shading indicating mean coral cover averaged over 1985–2012. Points show the locations of the 214 survey reefs in the northern, central, and southern regions, and their color indicates the direction of change in cover over time. (B) Box plots indicate the percentiles (25%, 50%, and 75%) of the coral cover distributions within each year and suggest a substantial decline in coral cover over the 27 y” (De’ath, 2012)

For the whole Great Barrier Reef, the mean coral cover has declined non-linearly from 28.0% to 13.8%, declining a total of 14.2% over the 27 years. This is equivalent to a loss of 50.7% of the original cover. Moreover, two-thirds of the decline has been since 1998, and decline rates increased to become consistently over 1.4% per year since 2006. In 2012, the last year of the study, the rate had risen to 1.51% over that year.

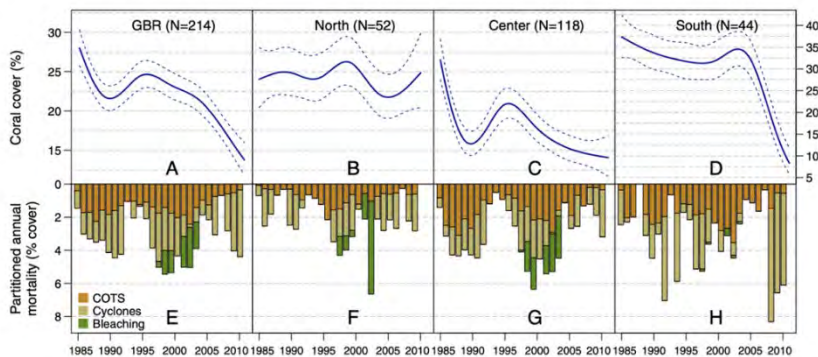


Figure 4. “Temporal trends in coral cover (A–D) and annual mortality due to COTS (Crown of Thorns Starfish predation), cyclones, and bleaching (E–H) for the whole GBR and the northern, central, and southern regions over the period 1985–2012 (N, number of reefs). (A–D) Trends in coral cover, with blue lines indicating estimated means (± 2 SEs) of each trend. (E–H) Composite bars indicate the estimated mean coral mortality for each year, and the sub-bars indicate the relative mortality due to COTS, cyclones, and bleaching. The periods of decline of coral cover in A–D reflect the high losses shown in E–H.” (De’ath 2012)

As seen in Figure 4, the northern reefs have had a relatively small change in mean coral cover over the 27 years, whereas the central and, more recently, southern reefs have declined significantly. The study also found that acute disturbances, namely Crown of Thorns Starfish predation (COTS), cyclones, and bleaching, had affected all but three of the 214 reefs. For the Torres Strait Islands, situated north of Cape York (which is at the northern tip of Queensland), only the northern reefs are of relevance. A more up to date study is a 90-year assessment of the well-documented Low Isle reefs of the northern Great Barrier Reef (Fine et al., 2019). According to this examination, it was found that the reefs of the Low Isles displayed significant visual deterioration in both soft and hard coral density from 1928 to 2019 with little sign of adaptation. Additionally, the study found the level of species richness in 2019 to be only 37% of that initially recorded in 1928.

Examining the three principal disturbances— COTS, cyclones, and bleaching events—is crucial to understanding the current and future outlook of the reef. As seen in Figure 3, a reducing population of COTS seemed to be evident in the north. This is primarily attributed to improved water quality (Schaffelke and Anthony, 2015). The invasive species are especially damaging to reefs, and thus the reduction in numbers is encouraging.

However, less optimistic is the outlook concerning coral bleaching events. Coral bleaching describes the phenomenon by which coral exoskeletons release their zooxanthellae (a symbiotic dinoflagellate), which provide nutrients to the coral via photosynthesis. The name arises from the changed appearance of corals from colorful to being white as if they were bleached. The widespread consensus is that sea temperature increases and exposure to higher solar radiation (albeit to a lesser extent) are the main factors that induce such reactions of corals (Brown, 1997). The increase of sea surface temperatures affects most severely the northern reefs, which are closer to the equator and thus in warmer waters (Australian Institute of Marine Science, n.d.). The loss of photosynthesizing zooxanthellae results in many corals dying of “starvation” after prolonged exposure to the altered environmental conditions (Douglas, 2003).

Over the last two decades, the frequency and severity of mass bleaching events have dramatically risen, and the Great Barrier Reef has suffered five devastating bleaching events (in 1998, 2002, 2016, 2017 and 2020) (Readfearn, 2020). What is absent from the 2012 report by De’ath et al. is the effects of the post 2016 bleaching events, which were significantly more devastating than those from previous coral bleaching events. For instance, the 2016 bleaching caused the northern stretch of the reef to suffer a further 50.3% loss of coral cover in eight months (Hughes et al., 2018). Research suggests that these events will only increase in frequency and geographical extent as average temperatures and thus average sea temperatures continue to rise (Sully et al., 2019).

The third main contributor to reef destruction is cyclones. Cyclones have great destructive power over coral reefs, which act as a natural barrier to soften the effects of tropical storms (ARC Centre of Excellence in Coral Reef Studies, 2018). In the past decade, several category 5 cyclones have hit the region, increasing coral mortality (Dowdy, 2021). Despite tropical storms increasing in severity and frequency in many areas globally, they are not projected to increase in frequency, but only in intensity, in Australia (Great Barrier Reef Marine Park Authority, n.d.).

Although not being a reason for coral cover loss, calcification rates are

also essential to coral reef health. The calcification rate supports reef recovery and further coral growth, making it a key indicator of the current health and future health of the ecosystem (United States Environmental Protection Agency, n.d.). However, multiple sources reported a declining rate of calcification in the Great Barrier Reef, which is a threat to the recovery of the ecosystem (De'ath, 2009). One cause for the decline is ocean acidification. Ocean acidification occurs when CO₂ is dissolved into the seawater and increases the pH. Calcium carbonate structures such as coral exoskeletons, shells, and other skeletons of marine animals are prone to disintegration by the increasingly acidic conditions (Mollica et al., 2018). Consequently, there is a decreasing concentration of carbonate ions, which is crucial to the sustenance of sustainable calcification rates able to encourage replenishing of the reef (Hoegh-Guldberg et al., n.d.).

Species richness is the other component of assessing the health of the Great Barrier Reef. A 2018 report examined literature to find significant declines in the health of seagrass ecosystems in addition to coral ecosystems (Richards et al., 2018). Also declining was the population of dugongs, hawksbill turtles, large fish populations, and several species of sharks and rays. More than 75% of the over 1,600 species of fish that inhabit the reef have declined in abundance, and around 50% of species have populations equivalent to half their original numbers in 1995 (Gabbatiss, 2018). Even before the major bleaching events post 2016, in 2002 it was documented that the diversity of fish in the four focal families of the reef had fallen by around 22% since 1997 (Jones et al., 2004). More concerning is that the majority of the decline has been concentrated in the northern part of the reef—with larger fish migrating south and smaller fish remaining in a weakened state, prone to predation (Loomis, 2018). Reports across the topic show increasing numbers of endangered species and declining populations—including those of keystone species (Whitmore et al., 2021).

After a review of current literature, it is evident that there is extensive agreement about an excessive loss in coral cover and a continuance of a similar—potentially worse—trend in the future. Perhaps correlated, there is also a widespread and clear fall in species diversity and abundance across the reef.

2.2. Health of the Torres Strait Islanders

As with most Indigenous communities, the health of the Torres Strait Islanders is on a separate level from that of the Australian mainland (Lowitja Institute, 2021). In the last two decades, there has been an important rise in interest and research dedicated to understanding and following the health of the Islanders. This review will focus on the health disparity and the mental and physical wellbeing of the Islanders.

Improving the health status of Indigenous peoples in Australia has long been a priority, but the disparity remains concerningly large (Durey and Thompson, 2012). One study led by Vos in 2008 delved into the details of this disparity (Vos et al., 2008). The study utilized the burden of disease approach on national population health datasets and Indigenous-specific epidemiological studies. This approach involves the calculation of current rates of Disability-Adjusted Life Years (DALYs) for Indigenous peoples and that for non-Indigenous peoples in Australia. DALYs are the principal measure for overall burden of disease, considering years lost due to premature mortality and years

lost lived in sub-optimal health. The two are then compared to give adjusted DALYs for the Indigenous population if they had the same level of mortality and disability as those not from an Indigenous background. The report found that if, in 2003, Indigenous Australians had the same burden of disease as the total population, 59% of the total DALYs of Indigenous Australians, or 56,455 DALYs, would have been avoided. From the database, the study found that 70% of the health gap was attributed to non-communicable diseases—with cardiovascular disease being the most prevalent health issue, followed by diabetes, mental disorders, and chronic respiratory diseases. Additionally, the main risk factors accountable for health compromises were, in order of decreasing prevalence: tobacco use (17%), high body mass (16%), physical inactivity (12%), high blood cholesterol (7%) and alcohol consumption (4%). Around 26% of the Indigenous population subsists in remote areas, but they suffer a disproportionate fraction of the health gap (40%). Another aspect of concern is the young demographic structure of the Islanders, an average age 15.6 years younger than the general population (Australian Institute of Health and Welfare, 2017). The youthful population presents potential for a further widening of the health inequality in following decades as the demographic shifts.

The statistics may suggest that there is not much intervention from the government to support the closing of this gap; however, not only has the Australian government acknowledged this disparity, but it has also been producing “Close the Gap” reports annually since 2008. The reports analyze current health circumstances and propose strategies for future interventions. A 2007 paper published by The Australian Human Rights Commission named “Social Determinants and the Health of Indigenous Peoples in Australia – a Human Rights Based Approach” illustrates the disparity very well. Although there have been many health advancements made in recent years, the rates of improvements are still second to that for the general population in Australia. For instance, whilst the life expectancy increased faster for the Indigenous population in 2010-2012, the general population has seen a 70% decrease in cardiovascular disease death rates over the last 35 years whereas there was no apparent reduction in those for the Indigenous population. At first sight, it may seem illogical for there to still be such a large disparity, for there have been specifically directed interventions for almost a decade now; nonetheless, one soon realizes the numerous factors influencing this sizeable health gap.

Albeit a complex issue, the three principal factors driving the gap are the socio-economic inequality, cultural barriers, and historic and systemic discrimination. Despite improvements in the last few decades, the economic inequality suffered by Indigenous peoples is clear and significant (Australians Together, 2017). Compared to the non-Indigenous population, Indigenous Australians are more likely to rely on government pensions and benefits as their main source of income. The average household income was lower by around \$300 per week, and the proportion of 15- to 64-year-olds unemployed was approximately double that of non-Indigenous Australians (Australian Institute of Health and Welfare, 2017). Even though Australia has universal healthcare that is, for the most part, free, Indigenous peoples do not have equal access to primary healthcare and sufficient infrastructure (Harfield et al., 2016). For instance, there is insufficient infrastructure to provide safe drinking water, effective sewerage treatment, and healthy housing (Infrastructure Australia,

2012), although the Queensland government has pursued developing more healthcare facilities in remote areas. It now has facilities at 13 of the 18 Torres Strait Islands (Queensland Government, accessed 2021).

Economic disadvantage has also affected the nutrition of the Indigenous population (Sjöholm et al., 2019). In general, the young are more likely to be underweight, and the prevalence of overweight/obesity increases with age. However, in more remote and poor areas, there was a lesser proportion of people overweight, and a greater proportion underweight across all ages. It is apparent that socio-economic inequality has contributed to the health gap directly and indirectly.

Inequity in healthcare is also a consequence of cultural differences and barriers between the Indigenous population and the general population. It is crucial to note here that the Indigenous population consists of the Torres Strait Islanders and the Aborigines, who have differing cultures. However, the term “Indigenous culture” is used to refer to both cultures inclusively. In a thorough paper written by Li in 2017, some of the obstacles faced by Indigenous peoples include differing languages, medical practices, and perceptions of sexuality and gender. Such differing views often lead to serious miscommunication, which was the most common reported cause of unsatisfactory outcomes following interactions with healthcare services. An unwillingness to use mainstream health services was also documented, mainly because of differing beliefs and understanding about value, health, and identity. Common experiences also include delayed access to free health checks, cancer screening, and follow-up appointments, which exacerbate the hesitant attitude adopted towards mainstream healthcare services.

Perhaps the most prominent reason for the gap is the historical and systemic discrimination towards Indigenous peoples. The poverty and inequality they experience today are direct consequences of colonization and their historical treatment, and their continued hindrance to equal healthcare can be associated with systemic discrimination (Dick, 2007). Historical examples, such as the forcible removal of Aboriginal and Torres Strait Islander children from their families during the assimilation period, illustrate how even after several generations, Indigenous peoples are suffering from historical discriminatory treatment. Commonly referred to as the “stolen generation,” the National Aboriginal and Torres Strait Islander Social Survey 2002 reported that 38% of respondents had been removed forcibly from their natural families (Australian Bureau of Statistics, 2002). Such historical treatment has caused intergenerational health impacts and has undoubtedly decremented the level of trust between many families and the national healthcare services.

For the same reasons, there is not only a gap in physical health between Indigenous peoples and the general population, but also a stark mental health gap (CDC, 2018). Indigenous peoples globally continue to suffer from sub-optimal mental health, primarily due their identity, or rather their loss of identity, and historical treatment. Observed worldwide, PTSD-like symptoms are apparent in various Indigenous communities, with many experiencing inter-generational effects and consequences (American Psychiatric Association, n.d.). Different Indigenous peoples are at different positions in their journey to achieve equity. For instance, whilst the Natives of Alaska have achieved a high level of self-determination (Anderson, 2003), Indigenous people of Myanmar

are currently subject to human rights violations (ICCA Consortium, 2021). However, even when equity in legislation is achieved, it rarely equates to true equity in society—Indigenous groups are often subject to wide health disparities and socio-economic inequality (Belcourt, 2018). Having also experienced historical discrimination and trauma, they are essentially universally more vulnerable to mental health compromises. Such compromises often lead to lower self-esteem, a tendency for substance abuse, and unemployment (Louth, 2012). In response, guilt increases and self-esteem plummets, creating a negative cycle that is difficult to escape.

Acknowledgement of maltreatment such as genocide and of land ownership have been shown to be significant to the improvement of the self-esteem and overall mental health of Indigenous peoples (U.S. Department of Arts and Culture, n.d.). In Australia, land ownership was acknowledged in 2013, and a continuance of pushes for self-determination of Torres Strait Islanders was carried out by the Queensland government. However, despite legislation change, racism towards and trauma within Torres Strait Islander and Aboriginal communities are still commonplace (Markwick et al., 2019). These attitudes not only negatively impact the people's mental wellbeing, but also directly fuel a mental health disparity between the general population and Indigenous Australians. The Australian government claims that mental health inequality is responsible for 10% of the health gap and is funding mental health rehabilitation facilities specifically for Indigenous peoples (Australian Government Department of Health, 2021). Nonetheless, the hesitant and fatalist attitude many have towards the central healthcare system will be obstructive in the facilities' reaching their full potential.

Most significant is the intergenerational trauma evident in Australia's Indigenous communities. Intergenerational trauma exists with the survivors of governmental programs that had forcibly removed children from their natural families to indoctrinate European values and annihilate Indigenous traditions (Bintarsari, 2018). Aborigine children of parents who were involved in the programs were found to be twice as likely to develop behavioral and emotional problems, to be at high risk of hyperactivity, emotional and conduct disorders, and to be twice as likely to abuse drugs and alcohol (Western Australian Aboriginal Child Health Survey, 2002). Somewhat applicable to Torres Strait Islander children too, the same survey documented that 21.6% of Indigenous children under the age of 12 have experienced racism in the last 6 months. The survey also found a relationship between those who have experienced discrimination and tendencies of increased smoking, marijuana use and alcohol consumption in children under 12. The relationship between racism and mental health issues is observed universally, and often leads to substance abuse (Gibbons et al., 2010).

Another study reported a finding that 21% of Indigenous children (of a sample size of 3993 children) in Australia were at risk of developing clinically significant emotional or behavioral difficulties (CSEBD) (Shepherd et al., 2012). The study's results indicated that lower socio-economic statuses—which often impaired family functioning abilities, parenting quality, and living conditions (such as overcrowding) —played a crucial role to one's risk of CSEBD.

Of the overall Indigenous population, surveys agree on the elevated rates of psychological distress (depression and anxiety) experienced among the

communities, which were almost three times higher than that of the general Australian population (Cunningham and Paradies, 2012). Many studies also report the conceptions of mental illnesses as weaknesses and thus a related unwillingness to seek for help (Parker and Milroy, 2013). Such high rates are direct consequences of the large number of stressors the people encounter. A survey by the Australian Bureau of Statistics defined stressors to include “the death of a family member or close friend, serious illness or disability, inability to get work, overcrowding at home, and alcohol and drug-related problems” (n.p.). The results illustrated that 44% of the Indigenous population have suffered at least 3 stressors in the previous 12 months, and 12% suffered at least seven. Reportedly, traditional cultural practices, which relieved anxiety and stress, were vital in helping people tolerate their frequent exposure to multiple stressors (Trewin and Madden, 2005). Hunter and Eastwell noted in their research the tendencies of Torres Strait Islanders and Aboriginal peoples to express very strong feelings within specific cultural contexts such as funerals.

Over the year 2018-2019, a self-assessment survey for the total Indigenous population saw 24% of people report having a mental health condition; of these, 30% was due to anxiety and depression (Australian Institute of Health and Welfare, 2020). With the context of the cultural background, where there is a general reluctance to admit mental illnesses, a self-reported survey will most likely yield underestimations or a lower bound of true figures. A worrisome statistic is the suicide rate among Indigenous peoples. Persistently, they have suffered from significantly higher suicide rates. In 2021, suicide rates for those aged 10-19 were 4.1 times those for non-Indigenous Australians, and rates across every age group for Indigenous peoples were higher (Gibson et al., 2021). According to the Australian Institute of Health and Welfare, suicide rates nearly doubled from 2003 to 2019 for Indigenous peoples—from an age standardized rate per 100,000 persons of 14.5 to 27.1. For non-Indigenous people over the same period of time, an increased age standardized rate of 2.1 was seen (Australian Institute of Health and Welfare, 2021). Despite the source expressing caution due to “data quality issues,” including the under-identification of Indigenous peoples in deaths data and the uncertainties in population, it is safe to assume that there is a rising trend in suicide rates of Indigenous peoples.

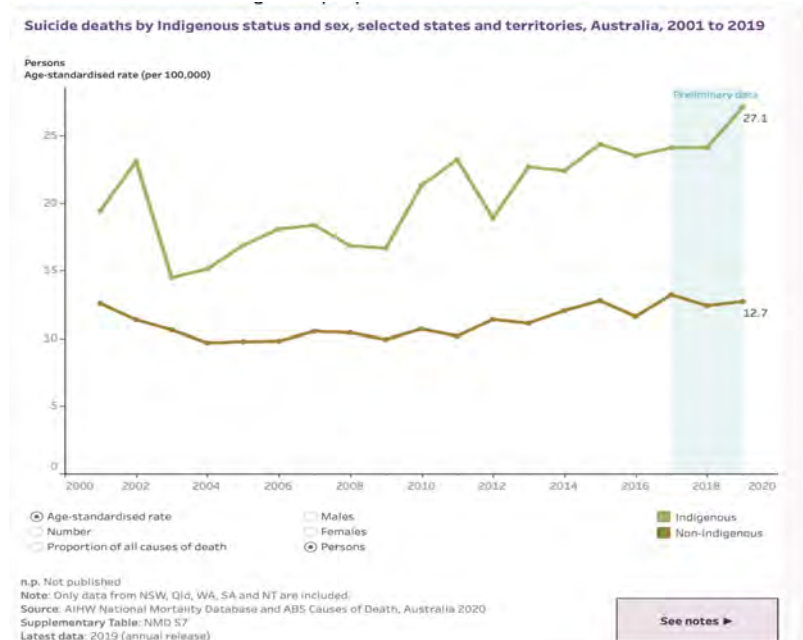


Figure 5. Suicide death rates by Indigenous Status in States of NSW, Queensland, West Australia, South Australia, and Northern Territory from 2001-2019. These are places where the majority of Indigenous (especially Torres Strait Islanders) reside and thus are a decent representation of the trend in suicide rates. (Australian Institute of Health and Welfare, 2021)

This brief overview of the physical and emotional wellbeing of the Indigenous population provides clear evidence that in both aspects, Indigenous peoples are at substantial disadvantages to the general population. Moreover, despite the efforts of the Australian government to fund infrastructure, do research, and modify legislation, there seems to be no trend of the gap closing (Fisher et al., 2021). The review also highlights dilemmas that exist with the attitudes towards centralized healthcare and illustrates that shifting this attitude will be a challenging obstacle that must be overcome if equity is to be achieved.

3. Discussion

3.1. Indigenous Communities' Reactions to Changes in their Environments

Environmental change is happening worldwide and is disproportionately affecting Indigenous communities (UN Department of Economic and Social Affairs, n.d.). In this context, climatic and environmental change represents an issue of socio-environmental injustice.

With a very limited number of resources available that focus on the relationship between the Torres Strait and the ever-changing Great Barrier Reef,

a review of several other Indigenous peoples in similar situations will be beneficial. One example is the Iñupiaq people, one of the Native peoples of Arctic Alaska. Although they have vastly different cultures and traditions to Torres Strait Islanders, the Iñupiat have also suffered tremendous losses in recent years due to climate change (Sakakibara, 2008). Sakakibara (2020) also describes the forced migration of the community from their spiritually and emotionally significant home, Point Hope, to set up a new Point Hope more inland due to life-threatening coastal erosion. Flooding has drowned their traditional locations of great significance, such as their cemeteries and ceremonial grounds, and permafrost and shores are melting away as the sea ice recedes. For the Iñupiat, one coping mechanism to this emotional loss is the act of storytelling. Practiced by Elders, storytelling provides a means to resuscitate a sense of home and a sense of place through their memories. More innovatively, a young educator of the community created digital versions of some of the lost sites that were culturally significant so that current and future generations are still able to “visit” and appreciate the sites. The analysis of this specific community of Alaska Natives demonstrates the turbulence climate change has caused and the resilience of the community—driving it to make the best of its situation.

Nonetheless, the research suggests that this loss due to environmental change has caused a significant negative impact to the mental health of the community. There was a widespread sense that their home was drowning and that their way of life was also slipping away. Anxiety and depression are more prevalent already in their society, even though there is evidence that methods of bringing people closer to their cultural background helped them overcome such difficulties. Therefore, it can be assumed that this loss of heritage and inability to visit traditional sites have had negative effects on the overall mental health of the Iñupiaq. Spiritually and emotionally, the people were devastated, but they were also affected physically. The Iñupiaq relied heavily on the catching of bowhead whales as communal celebrations and as a source of food. They are deeply connected with the species and even have a term for the complex entanglement of the human body and the migration route of the bowhead whales. Recently, climate change has influenced the migration route of the whales, and there is a general sense within the community that their ways of life and wellbeing are at risk. The Iñupiaq are very tightly connected spiritually and physically to their physical environment. Hence, it has been exceptionally challenging for them to accept and adapt to the dramatic changes to their home.

Another Indigenous community threatened in the face of climate change is the Swinomish Indian Tribal Community of Washington State (Centers for Disease Control and Prevention, 2020). The Swinomish Reservation located on Fidalgo Island is especially prone to the effects of climate change. Situated off the coast of Seattle, the island has a sizable saltwater coastline and a dangerously low average elevation. Hence, threats of increased rates of coastal erosion and sea level rising pose the greatest risk to the community. Traditionally, the Swinomish have always been a predominantly fishing tribe—relying on salmon, forage fish, marine mammals, shellfish, and wild game as sources of food and other resources (U.S. Geological Survey, 2015). Such organisms have become quintessential in their culture and play a major role in their traditions and community wellbeing, enhancing the people’s sense of identity. The approach taken by the Swinomish is to better adapt to changes in the environment to

preserve their identity. To this end, the Swinomish Tribe have implemented the Swinomish Climate Change Initiative dedicated to identifying the entire range of probable impacts climate change may induce. Investigating factors such as the rapidly declining shellfish population and the damage likely to affect infrastructure—especially culturally significant buildings, the initiative analyses the consequences and their potential impacts on the health of the community. Such assessments—ecological and societal—are then combined to establish tribal priorities and form future adaptation strategies (Morrison, 2020). The establishment of such a strategy may be comforting to the community by providing a feeling of preparedness for the changing way of life in the future.

Whilst there are many more Indigenous communities being forced to adapt to changing environments (United States Department of Agriculture, 2014), the Iñupiat and Swinomish peoples demonstrate the vulnerability of coastal Indigenous communities. Similar to the Iñupiat and Swinomish peoples, the Torres Strait Islanders are also deeply connected to their environment—with the Great Barrier Reef forming a major part of it. Thus, the culture, lifestyle, and identity of the Torres Strait Islanders are also at risk. Despite different Indigenous peoples having contrasting situations and thus having varying abilities to cope and adapt, all coastal Indigenous communities will experience additional hardship as their traditional lifestyle, culture, and identity are put at risk (UN Department of Economic and Social Affairs, n.d.). Inevitably, climate change will impact the peoples' mental health and the overall community's wellbeing—pressuring communities to adapt in an attempt to preserve their traditions and ensure the continuation of their identity.

3.2. The Torres Strait Islanders

The paper has briefly covered some important roles the Great Barrier Reef plays for the Torres Strait Islander communities. However, it is crucial to evaluate their dependence on the surrounding ecosystem with more detail. Traditionally, the Torres Strait Islanders have always been agriculturalists who have relied primarily on hunting and gathering for food (Torres Strait Island Regional Council, n.d.). Cultural foods commonly included pork, local fruits, and vegetables such as yam and cassava, and marine animals including dugongs, fish, shellfish, stingrays, and turtles (Queensland Government, n.d.). Cultural practices such as distinct methods of cooking and hunting are taught to children at a young age, and others—such as dancing celebrations, music, and intergenerational storytelling—are intertwined into communal life (Torres Strait Island Regional Council, n.d.). Recognizing the Islanders' dependence on the Great Barrier Reef, the government has permitted their traditional uses such as fishing, collecting, and hunting, despite the Great Barrier Reef being a highly restricted Marine Conservation Park.

The culture of Torres Strait Islanders is multi-layered and incredibly complex so only several relevant aspects will be described. To the Islanders, the gathering, preparation, and consumption of meals have always played a key part in their lives (Queensland Health, n.d.). Eating or sharing food sources within a community is a social event, and the importance of the event is evident by the many rules that govern which foods are forbidden for consumption. Not only does this ensure a sustainable lifestyle, but it also reinforces the peoples' identity

for, depending on various factors, each person and each clan has a totem, which is an animal they are banned from harming or eating (Queensland Health, n.d.). Such traditions were facilitated by a deep knowledge of the surrounding environment, including the analysis of the wind, tides, and rain, to determine the change of seasons and prey (Queensland Health, n.d.). Sources of food from the sea depends on the season, but most commonly involves dugongs, turtles, stingrays, fish, and shellfish. Many of these animals also possess significant roles in traditional ceremonies and celebrations, adding to the community's wellbeing and identity (Queensland Health, n.d.). Thus, with their decline in numbers, the Islanders have had to adapt their traditional diet, ceremonies, and feasts to ensure that the population of such animals is not further impacted. Additionally, with some species which are severely endangered, such as several of the turtle species, the Islanders have had to sacrifice their traditional lifestyles. As seen with other Indigenous populations, the sacrifice of consuming customary foods and ways can be especially detrimental to the mental health and wellbeing of the people (Skinner and Pratley, 2016).

A variety of reasons have already been discussed for the existence of the mental and physical disparity between Indigenous Australians and the general population. Such factors included socio-economic inequality, cultural differences, and racism; but a potential factor not discussed is the loss of tradition.

To Indigenous communities worldwide, tradition has always played a principal role in supporting the mental and physical wellbeing of the people. It is no different with Torres Strait Islanders. A study published in 2021 reports lower suicide rates among those with higher levels of cultural connectedness, which consisted mainly of participating in cultural events, and higher suicide rates among those in areas with a greater reported prevalence of discrimination (Gibson et al., 2021). Thus, with a decline in traditional foods, there is a risk of the people's mental health degrading as related culturally-significant practices are negatively affected—especially with the significance of meals to the Islanders.

Nonetheless, the risks to the Torres Strait Islanders' culture may not solely be limited to a loss of their traditional lifestyle. Climate change is also threatening their ability to continue living on their traditional land. Various consequences of climate change—including high tides, strong unpredictable winds, coastal erosion, and rising sea levels—are also affecting the Islands. In 2012, the island of Saibai suffered from severe flooding, which devastated several of the sacred sites, including the cemetery. In 2017, flooding on a different island (Boigu Island) washed away sections of road, and in 2018, Yam Island had floods that left many homeless and dispossessed (Lopez, 2019). The rising sea levels have forced many to migrate as communities struggle to find non-swampy ground with uncontaminated groundwater and soil. Leaving their community, which was built upon the surrounding ecosystem, is undoubtedly difficult, and prompts the experience of solastalgia. As mentioned before, solastalgia describes “the pain or sickness caused by the loss or lack of solace and the sense of isolation connected to the present state of one's home and territory” (Albrecht et al., 2007, p. 45). In the paper defining solastalgia by Albrecht et al., there is a quotation by E. Mitchell (1946), who asserts, “Divorced from his roots, man loses his psychic stability.” Seeking refuge in areas away

from the Torres Strait Islands, where there is a lower threat from the effects of climate change, will unquestionably cause a disconnection from their land to which they are spiritually and physically attached. The land represents their roots and has played a key part to their history. Thus, this forced separation and abandonment of their land will affect the mental wellbeing of the Islanders greatly. Solastalgia commonly causes widespread distress and can often escalate to induce serious health issues such as substance abuse and mental illnesses such as depression and suicide (Albrecht et al., 2007). Also noted in the study is the likelihood of a community distressed by solastalgia manifesting an array of social problems, such as overly high rates of crime, which often leads to “community dysfunction and crisis” (Albrecht et al., 2007, p. 47). The disconnection will incite a reduced sense of belonging, a greater risk of facing discrimination, and a sense that they are deserting their ancestry. Cultural events will be reduced even further, and communities may become split post-migration. To many, it will be a slow and painful process to see their native communities and tradition slowly disintegrate.

Moreover, the loss of land to the environment may be a reminder of the colonization experience, where land was taken away by European settlers. Clearly portrayed by the annual celebrations the Islanders enjoy on the day they obtained ownership rights to their land (June 3rd), the possession of their land is crucial to the community's wellbeing. Thus, the losing of their land once again may trigger distress across the peoples and resurrect a feeling of displacement, unworthiness, and unbelonging.

A lesser form of the forced migratory process started several years ago, so it is also a potential factor contributing to the apparent decline in the mental health of the community. However, future projections suggest the exacerbation of this trend as the climate continues to change and consequences increase in frequency and effect (IPCC, 2021). This may further the mental health disparity between Indigenous Australians and non-Indigenous Australians, and like a chain reaction, it may elevate rates of substance abuse and advance the disparity in physical health too. It may also negatively impact the socio-economic status of the peoples as they are dispossessed, displaced, and forced to settle in a new environment.

3.3. Solutions

The future outlook may appear bleak. Nevertheless, there are several advancements underway which seek to reduce the severity of consequences. The first step is to recognize the voices and experiences of historically marginalized and underrepresented populations of Indigenous peoples including the Torres Strait Islanders. On February 13, 2008, then Prime Minister Kevin Rudd made a formal apology on behalf of the Australian Parliament to Australian Aboriginal and Torres Strait Islander peoples, in particular to the Stolen Generations. The national government and administration will have to further develop their effort to cultivate ethical partnerships with the Indigenous communities to address myriad issues associated with environmental and social injustice as well as solastalgia.

Furthermore, the primary development will be the global decarbonization and transition to using more green energy. Many factors

affecting the communities are derived from climate change, and thus the pursuit of reducing the amount of change will be crucial.

Other potential solutions which are at various points of development include coral shading, assisted breeding, manual removal, water quality improvement and genetic engineering (Anthony et al., 2020). Coral shading describes the artificial manipulation of clouds to cool the region underneath by inducing rain or by reducing solar radiation. Additionally, this could also be done by applying thin films on the surface of the water, using algae, or misting the area above. The technology for coral shading, being relatively novel, is still in its early stages of development.

Assisted breeding involves in vitro fertilization (IVF) attempts and other methods of trying to breed corals and grow them (Great Barrier Reef Foundation, 2020). Because of a low frequency and a short duration of spawning seasons, coral eggs and sperm are collected and fertilized artificially. The corals are then grown under surveillance until big enough to reintroduce to the wild. Results from the IVF trial in 2016 shows promising results as manually bred corals are seen to be thriving. This would be crucial in restoring the ecosystem around the Torres Strait Islands and recovering surrounding biodiversity.

Manual removal programs and water quality improvement programs are both targeted at reducing the number of Crown of Thorns Starfish (COTS), which directly threaten the health of reefs. The invasive species has been documented to thrive in areas with greater sediment concentration and waters with more pollution (Wescott et al., 2020). Manual removal programs take time and effort but have shown the most evident results in diminishing the population of the starfish (Wescott et al., 2020). Compared to the other potential solutions, these two programs are well established and do not involve novel technology. However, with the advancement of AI, it could be possible one day to have robots that can recognize and remove COTS.

As some coral reefs are situated in warmer regions than others, there are some reefs that have been able to withstand the increased ocean temperatures. Some mutations have also led populations of corals to obtain the ability to tolerate a greater range of temperatures. Such resilient corals may be studied and used to genetically engineer other corals which are also able to tolerate a greater range of temperatures. With increasing knowledge of breeding and growing corals, these genetically modified corals may then be supported until ready to be reintroduced into the wild. Subsequently, descendants of such corals would also have such resilient characteristics and could increase the overall preparedness and tolerance of coral reefs for rising temperatures.

Certainly, there is a lot of scientific interest and public support for the conservation of marine ecosystems. With so many developing technologies, perhaps some will be able to reduce the impacts of climate change and ecosystem degradation. The Torres Strait Islanders are particularly vulnerable to the consequences of climate change, and the slowdown of climate change will provide time to adapt so that they suffer a less severe loss of culture and heritage.

A viable solution would be the establishment of dedicated communal sites on the mainland. A coordinated and financially supportive program to relocate Islanders living in areas of risk allows them to settle and maintain a level of cohesion. The development of cultural sites around the regions would

also be beneficial to the relocation, as it will increase their sense of belonging and thus their mental health. Admittedly, their lifestyles will be inevitably altered and challenged, but their culture and community can continue to thrive with the appropriate support. Thus, programs and plans developed in advance to help the communities transition are vital for the sustenance of the Torres Strait Islanders.

4. Conclusions

It is clear that the reef as a whole is deteriorating—rapidly losing both its percentage of coral cover and biodiversity. However, studies show that different regions of the natural structure are degrading at different rates. With the increase in frequency and damage of coral bleaching events, coral mortality rates throughout the reef are rising. In particular, northern reefs—which are where the Torres Strait Islands are situated—are experiencing higher bleaching and mortality rates as a result of the reef's close proximity to the equator. Additionally, increases in ocean acidity, COTS population, and the severity of tropical storms, are also factors threatening the future prosperity of the reef. In terms of biodiversity, sources reveal the apparent migration of species away from the northern regions into the cooler and less damaged regions in the South. As waters continue to get warmer, populations and species richness will continue to decline in the North.

It must be noted that although the resources focused on the health of the Torres Strait Islanders is increasing, they are still relatively limited. Evidently, the Islanders' general health and accessibility to primary healthcare has improved significantly in the last few decades. However, there is still a large health disparity between the general Australian population and the Indigenous Australian population, despite having a national program committed to closing this gap since 2008. The evident hesitancy for Indigenous Australians to use the national health services is a challenging obstacle that must be overcome to close this gap. In addition to several reasons cited for the health disadvantage experienced by Indigenous Australians, this study proposes the loss of tradition as another major factor.

The Great Barrier Reef has always played a central role in the tradition and culture of the Torres Strait Islanders. Being what their lifestyles revolved around, and spiritually and emotionally connecting them to the surrounding biome, the declining populations of species that have major cultural significance threatens the continuance of many traditional events. Additionally, with meals so highly regarded in the Islanders' culture, it is likely that the reduction in the consumption of culturally important animals will diminish their sense of community, sense of family, and connection to legacy.

Furthermore, the changes to the Great Barrier Reef will also indirectly affect and cause distressing consequences to the Islanders. Surrounding coral reefs have always acted as a natural barrier to soften the impact of cyclones. Over the last decade, several islands have had to evacuate the island's people and permanently displace them due to concerns of further flooding and other weather events. Forced migration can be especially painful to Indigenous communities. One reason is that they are often split up post migration based on

socio-economic and other factors. It is hurtful to lose connection to their ancestral lands to which they are deeply connected. Because of the importance of cultural connection to the self-esteem and thus mental health of Indigenous peoples, it is likely that the communities will suffer from suboptimal mental health if their cultural connection plummets. Thus, with trends that suggest more Islanders will have to be displaced over the next decade, the government and public must take action to ensure the wellbeing of the Islanders.

Decarbonization of sectors such as energy production, transport, and industry are crucial to slow down the rate and damage of climate change, which Indigenous communities are most vulnerable to. However, direct intervention to help the local area will also be necessary. This involves the development and usage of new methods—such as coral shading, removal of invasive species, assisted breeding, and genetic engineering—to restore coral cover and biodiversity. Additionally, it will also be vital for the government to draw up plans that will ensure a smooth relocation for the Islanders to prevent a crushing sense of solastalgia.

To maintain a sense of community and cultural connection, Islanders should be encouraged to settle down in areas together, with designated infrastructure implemented to continue supporting their traditions. Their history and culture should also be taught more extensively in school to reduce the amount of potential racism they experience. Moreover, there should also be programs that will support them to find jobs in these new environments, as many will struggle to find jobs in the changed landscape. Other than governmental support, they should be accepted and welcomed by the general Australian population, especially since evidence shows the devastating effects of discrimination on people's mental health.

The overarching purpose of this study was to evaluate if the deterioration of the Great Barrier Reef has had any effect on the Torres Strait Islanders' health and wellbeing. The evidence shows the degradation of their natural environment has had a significant impact on the communities. Moreover, as the rate of deterioration is projected to worsen, further decline of the Great Barrier Reef will result in devastating consequences for the Islanders. Around the world, many Indigenous groups are also subject to threats presented by climate change, and it is my hope that this study inspires similar evaluations of the effects of the changing environment on the Native peoples. The goal is to improve the wellbeing and quality of life for all, and such studies provide insight, highlight concerns, and offer suggestions so that governmental agencies can act to ensure a smooth adaptation process for the native communities. Thus, the significance of this study is not only the conclusions drawn regarding the relationship between the health of the Torres Strait Islanders, but also the attempt to inspire more exploration into similar relationships globally.

References

- AIHW, Australian Institute of Health and Welfare. Australia's Welfare 2017: In Brief, Indigenous Australians, Australian Institute of Health and Welfare, 2017, www.aihw.gov.au/reports/australias-welfare/australias-welfare-2017-in-brief/contents/indigenous-australians. Accessed 11 Aug 2021
- AIHW, Australian Institute of Health and Welfare. Income and Employment for Indigenous Australians, Australian Institute of Health and Welfare, 2017, www.aihw.gov.au/getmedia/2f327206-c315-43a7-b666-4fe24fetc12f/aihw-australias-welfare-2017-chapter7-5.pdf.aspx. Accessed 11 Aug 2021
- AIHW, Australian Institute of Health and Welfare. "Suicide & Indigenous Australians." Deaths by Suicide amongst Indigenous Australians, Australian Institute of Health and Welfare, 2021, www.aihw.gov.au/suicide-self-harm-monitoring/data/populations-age-groups/suicide-indigenous-australians. Accessed 11 Aug 2021
- Albrecht, Glenn, Nick Higginsbotham, Gina Satore, and Brian Kelly. 2005. "Solastalgia: The Distress Caused by Environmental Change." *PAN/Australian Psychiatry*, 41-55. Accessed 11 Aug 2021
- American Psychiatric Association. "Stress & Trauma." Treating Indigenous People in a Changing Political and Social Environment, American Psychiatric Association, n.d., www.psychiatry.org/psychiatrists/cultural-competency/education/stress-and-trauma/indigenous-people. Accessed 11 Aug 2021
- Anderson, Trudy. "Alaska Native Health Care: A Profile of Successful Self-Determination." *Alaska Native Health Care: A Profile of Successful Self-Determination*, Cultural Survival, 1 Sept. 2003, www.culturalsurvival.org/publications/cultural-survival-quarterly/alaska-native-health-care-profile-successful-self. Accessed 11 Aug 2021
- Anthony KRN, Helmstedt KJ, Bay LK, Fidelman P, Hussey KE, Lundgren P, et al. (2020) Interventions to help coral reefs under global change—A complex decision challenge. *PLoS ONE* 15(8): e0236399. <https://doi.org/10.1371/journal.pone.0236399>. Accessed 14 August 2021
- ARC Centre of Excellence in Coral Reef Studies. "Coral reefs protect coasts from severe storms." *ScienceDaily*. ScienceDaily, 3 April 2018. www.sciencedaily.com/releases/2018/04/180403111123.htm>. Accessed 11 Aug 2021
- Australian Bureau of Statistics, National Aboriginal and Torres Strait Islander Social Survey 2002, ABS cat. no. 4714.0, Commonwealth of Australia, Canberra, 2004, pp5-6. About 8% of Indigenous respondents reported that they themselves had been removed from their natural family. The most frequently reported relatives removed were grandparents (15%), aunts or uncles (11%), and parents (9%). Accessed 11 Aug 2021
- Australian Census 2016. "Estimates of Aboriginal and Torres Strait Islander AUSTRALIANS, June 2016." Australian Bureau of Statistics, Commonwealth of Australia, 2016, www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/estimates-aboriginal-and-torres-strait-islander-australians/latest-release. Accessed 11 Aug 2021

- Australian Government Department of Health. "Aboriginal and Torres Strait Islander Mental Health Program." Australian Government Department of Health, 29 June 2021, www.health.gov.au/initiatives-and-programs/aboriginal-and-torres-strait-islander-mental-health-program. Accessed 8 Aug 2021
- Australian Government, AIMS. Coral Bleaching Events, Australian Institute of Marine Science, n.d., www.aims.gov.au/docs/research/climate-change/coral-bleaching/bleaching-events.html. Accessed 11 Aug 2021
- Australian Government, GBRMPA. "Storms and Cyclones." Storms and Cyclones, Great Barrier Reef Marine Park Authority, n.d., www.gbrmpa.gov.au/our-work/threats-to-the-reef/climate-change/storms-and-cyclones. Accessed 8 Aug 2021
- Australian Institute of Aboriginal and Torres Strait Islander Studies. "Torres Strait Islander Flag." AIATSIS, Australian Institute of Aboriginal and Torres Strait Islander Studies, aiatsis.gov.au/explore/torres-strait-islander-flag. Accessed 6 Aug 2021
- Australian Institute of Health and Welfare. Indigenous Health and Wellbeing, Australian Institute of Health and Welfare, 2020, www.aihw.gov.au/reports/australias-health/indigenous-health-and-wellbeing. Accessed 4 Aug 2021
- Australians Together. Indigenous Disadvantage in Australia, Australians Together, 2017, australianstogether.org.au/discover/the-wound/indigenous-disadvantage-in-australia. Accessed 11 Aug 2021
- Belcourt, Annie. The Hidden Health Inequalities That Indigenous Peoples Face, High Country News – Know the West, 26 Jan. 2018, www.hcn.org/articles/tribal-affairs-the-hidden-health-inequalities-that-indigenous-peoples-face. Accessed 11 Aug 2021
- Bintarsari, Nuriyeni Kartika. "The Cultural Genocide in Australia: A Case Study of the Forced Removal of Aborigine Children from 1912-1962." SHS Web of Conferences, EDP Sciences, 14 Nov. 2018, www.shs-conferences.org/articles/shsconf/abs/2018/15/shsconf_icolgas2018_05002/shsconf_icolgas2018_05002.html. Accessed 11 Aug 2021
- Britannica, The Editors of Encyclopaedia. "Torres Strait Islander peoples". Encyclopedia Britannica, 20 Oct. 2020, <https://www.britannica.com/topic/Torres-Strait-Islander-people>. Accessed 5 August 2021
- Brodie, Jon, and Richard G. Pearson. "Ecosystem Health of the Great Barrier Reef: Time for Effective Management Action Based on Evidence." Estuarine, Coastal and Shelf Science, Academic Press, 10 May 2016, www.sciencedirect.com/science/article/pii/S0272771416301469. Accessed 5 August 2021
- Brown, B.E. Coral Bleaching: Causes and Consequences, Springer, 1997, www.coralations.org/Brown%201997%20Blea%20causes%20conseq.pdf. Accessed 5 August 2021
- CDC, Centers for Disease Control and Prevention. "Swinomish Tribe Assesses Climate Change Impact." Swinomish Indian Tribal Community, Centers for Disease Control and Prevention, 3 Sept. 2020, www.cdc.gov/publichealthgateway/field-notes/2020/swinomish-climate.html. Accessed 5 August 2021

- CDC. (2018). Health United States, 2017. Table 46. <http://www.cdc.gov/nchs/data/hus/hus17.pdf>. Accessed 5 August 2021
- Close the Gap Campaign Steering Committee, Lowitja Institute. "Close the Gap (2021)." CLOSE THE GAP CAMPAIGN REPORT 2021, Australian Human Rights Commission, 2021, humanrights.gov.au/our-work/aboriginal-and-torres-strait-islander-social-justice/publications/close-gap-2021. Accessed 11 August 2021
- Commonwealth of Australia 2017, CAPAD 2016. File: Great Barrier Reef Marine Park Locator Map.svg, Wikimedia Commons, 2017, commons.wikimedia.org/wiki/File:Great_Barrier_Reef_Marine_Park_locator_map.svg. Accessed 14 August 2021
- Commonwealth of Australia. Aboriginal and Torres Strait Islander Heritage Strategy for the Great Barrier Reef Marine Park, Great Barrier Reef Marine Park Authority, 2019, elibrary.gbrmpa.gov.au/jspui/bitstream/11017/3425/7/GBRMPA_ATSI_HeritageStrategy.pdf. Accessed 11 August 2021
- Cunningham, J., Paradies, Y.C. Socio-demographic factors and psychological distress in Indigenous and non-Indigenous Australian adults aged 18-64 years: analysis of national survey data. *BMC Public Health* 12, 95 (2012). <https://doi.org/10.1186/1471-2458-12-95>. Accessed 13 August 2021
- Davy, C., Harfield, S., McArthur, A. et al. Access to primary health care services for Indigenous peoples: A framework synthesis. *Int J Equity Health* 15, 163 (2016). <https://doi.org/10.1186/s12939-016-0450-5>. Accessed 11 August 2021
- De'ath, Glenn et al. "Declining coral calcification on the Great Barrier Reef." *Science (New York, N.Y.)* vol. 323,5910 (2009): 116-9. doi:10.1126/science.1165283. Accessed 5 August 2021
- De'ath, Glenn et al. "The 27-year decline of coral cover on the Great Barrier Reef and its causes." *Proceedings of the National Academy of Sciences of the United States of America* vol. 109,44 (2012): 17995-9. doi:10.1073/pnas.1208909109. Accessed 3 August 2021
- Deloitte Access Economics, 2017. At What Price? The Economic, Social and Icon Value of the Great Barrier Reef. Brisbane. Accessed 11 August 2021
- Department of Economic and Social Affairs, UN. "Climate Change for Indigenous Peoples." *The Effects of Climate Change on Indigenous Peoples*, United Nations, n.d., www.un.org/development/desa/indigenouspeoples/climate-change.html. Accessed 1 August 2021
- Dick, Darren. Social Determinants and the Health of Indigenous Peoples in Australia – a Human Rights Based Approach, The Australian Human Rights Commission, 2007, <https://humanrights.gov.au/about/news/speeches/social-determinants-and-health-indigenous-peoples-australia-human-rights-based>. Accessed 8 August 2021
- Douglas, A.E. "Coral Bleaching—How and Why?" *Marine Pollution Bulletin*, Pergamon, 20 Mar. 2003, www.sciencedirect.com/science/article/abs/pii/S0025326X03000377. Accessed 5 August 2021
- Dowdy, Andrew. Tropical Cyclones and Climate Change in the Great Barrier Reef Region, Australian Government's National Environmental Science Program, 2021, nеспclimate.com.au/wp-content/uploads/2021/04/ESCC_Tropical-cyclones-in-the-GBR_Brochure.pdf. Accessed 14 August 2021

- Durey, A., Thompson, S.C. Reducing the health disparities of Indigenous Australians: time to change focus. *BMC Health Serv Res* 12, 151 (2012). <https://doi.org/10.1186/1472-6963-12-151>. Accessed 11 August 2021
- Fine, M., Hoegh-Guldberg, O., Meroz-Fine, E. et al. Ecological changes over 90 years at Low Isles on the Great Barrier Reef. *Nat Commun* 10, 4409 (2019). <https://doi.org/10.1038/s41467-019-12431-y>. Accessed 4 August 2021
- Fisher, Matthew, et al. "Stakeholder perceptions of policy implementation for Indigenous health and cultural safety: A study of Australia's 'Closing the Gap' policies." *Australian Journal of Public Administration*, vol. 80, no. 2, 2021, pp. 239-260. OhioLINK Electronic Journal Center, doi:10.1111/1467-8500.12482. Accessed 6 August 2021
- Gabbatiss, J. (2018). Fish populations in Great Barrier Reef collapse after devastating coral bleaching events. *The Independent*. Retrieved from <https://www.independent.co.uk/environment/fish-populations-collapse-great-barrier-reef-coral-bleaching-australia-a8294751.html>. Accessed 14 August 2021
- Gibbons, Frederick X et al. "Exploring the link between racial discrimination and substance use: what mediates? What buffers?" *Journal of Personality and Social Psychology* vol. 99,5 (2010): 785-801. doi:10.1037/a0019880. Accessed 11 August 2021
- Gibson, Mandy, et al. "Suicide Rates for Young Aboriginal and Torres Strait ISLANDER People: The Influence of Community Level Cultural Connectedness." *The Medical Journal of Australia*, 7 June 2021, www.mja.com.au/journal/2021/214/11/suicide-rates-young-aboriginal-and-torres-strait-islander-people-influence. Accessed 14 August
- Morrison, Jim. "An Ancient People with a Modern Climate Plan." *The Washington Post*, WP Company, 24 Nov. 2020, www.washingtonpost.com/climate-solutions/2020/11/24/native-americans-climate-change-swinomish/. Accessed 9 August 2021
- Great Barrier Reef Foundation. "First Coral IVF Babies on the Reef Thriving." Great Barrier Reef Foundation, 2020, www.barrierreef.org/news/news/first-coral-ivf-babies-on-the-reef-thriving. Accessed 14 August
- Hamacher, DW, et al. "Dancing with the Stars': Astronomy and Music in the Torres Strait." Sophia Centre Press, 2017, arxiv.org/pdf/1605.08507.pdf. Accessed 6 August 2021
- Hoegh-Guldberg, O., et al. "Impacts of 1.5°C Global Warming on Natural and Human Systems." IPCC Secretariat, 1 Jan. 1970, <https://helda.helsinki.fi/handle/10138/311749>. Accessed 12 August 2021.
- Hoegh-Guldberg O, Poloczanska ES, Skirving W and Dove S (2017) Coral Reef Ecosystems under Climate Change and Ocean Acidification. *Front. Mar. Sci.* 4:158. doi: 10.3389/fmars.2017.00158. Accessed 18 August 2021
- Hoegh-Guldberg, Ove, et al. "People and the Changing Nature of Coral Reefs." *Regional Studies in Marine Science*, Elsevier, 3 June 2019, www.sciencedirect.com/science/article/pii/S2352485518306637. Accessed 11 August 2021
- Hopley, D., Smithers, S., & Parnell, K. (2007). Foundations of the Great Barrier Reef. In *The Geomorphology of the Great Barrier Reef: Development, Diversity and Change* (pp. 18-41). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511535543.003. Accessed 11 August 2021

- Hourigan, T. (2014). "A strategic approach to address fisheries impacts on deep-sea coral ecosystems," in *Interrelationships Between Corals and Fisheries*, ed. S. Bortoneed (Boca Raton, FL: CRS Press), 127–145. doi: 10.1201/b17159-9. Accessed 9 August 2021
- Hughes, T.P., Kerry, J.T., Baird, A.H. et al. Global warming transforms coral reef assemblages. *Nature* 556, 492–496 (2018). <https://doi.org/10.1038/s41586-018-0041-2>. Accessed 12 August 2021
- Hönisch, Bärbel, et al. "The Geological Record of Ocean Acidification." *Science*, American Association for the Advancement of Science, 2 Mar. 2012, [science.sciencemag.org/content/335/6072/1058.full?sid=36e46d6d-633c-4008-882b-e49c8a715a24](https://www.science.org/doi/10.1126/science.1219811). Accessed 9 August 2021
- ICCA, Territories of Life. Indigenous Peoples under New Attacks in Myanmar: An Open Letter and Call to Action, ICCA Consortium, 22 Apr. 2021, www.iccaconsortium.org/index.php/2021/04/05/myanmar-attacks-indigenous-peoples-open-letter-call-action/. Accessed 4 August 2021
- Infrastructure Australia. "Essential Indigenous Infrastructure." *Progress and Action*, Infrastructure Australia, 2012, www.infrastructureaustralia.gov.au/sites/default/files/201906/p195_iacoag_2012_chapter4_7_ws.pdf. Chapter 4, Pages 70-91. Accessed 9 August 2021
- IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. Cambridge University Press. In Press. Chapter 2. Accessed 14 August 2021
- Jones, G. P., McCormick, M. I., Srinivasan, M., & Eagle, J. V. (2004). Coral decline threatens fish biodiversity in marine reserves. *Proceedings of the National Academy of Sciences of the United States of America*. Retrieved from <https://www.pnas.org/content/101/21/8251>. Accessed 6 August 2021
- Korff, Jens. Torres Strait Islander Culture, Creative Spirits, 26 June 2021, www.creativespirits.info/aboriginalculture/people/torres-strait-islander-culture#how-is-their-culture-different-or-similar-to-that-of-mainland-aboriginal-people. Accessed 11 August 2021
- Li, Jie-Li. "Cultural Barriers Lead to INEQUITABLE Healthcare Access for Aboriginal Australians and Torres Strait Islanders." *Chinese Nursing Research*, Elsevier, 17 Oct. 2017, www.sciencedirect.com/science/article/pii/S2095771817301044. Accessed 9 August 2021
- Loomis, I (2018). Heat Waves, More Than Coral Death, May Cause Fish to Flee Reefs. *EOS News by American Geophysical Union*. Retrieved from <https://eos.org/articles/heat-waves-more-than-coral-death-may-cause-fish-to-flee-reefs>. Accessed 18 August 2021
- Lopez, Jaime. Australia: As the Great Barrier Reef Shrinks, Torres Strait Islanders Have Everything to Lose, Minority Rights Group, 2019, [minorityrights.org/trends2019/australia/](https://www.minorityrights.org/trends2019/australia/). Accessed 14 August 2021

- Markwick, A., Ansari, Z., Clinch, D. et al. Experiences of racism among Aboriginal and Torres Strait Islander adults living in the Australian state of Victoria: a cross-sectional population-based study. *BMC Public Health* 19, 309 (2019). <https://doi.org/10.1186/s12889-019-6614-7>. Accessed 12 August 2021
- Mollica, Nathaniel R., et al. "Ocean Acidification Affects Coral Growth by Reducing Skeletal Density." *PNAS*, National Academy of Sciences, 20 Feb. 2018, www.pnas.org/content/115/8/1754. Accessed 9 August 2021
- Mullins, S. *Torres Strait: A History of Colonial Occupation and Culture Contact, 1864-1897*. 1, CQUniversity, 1 Jan. 1995, <https://hdl.handle.net/10018/34271>. Accessed 4 August 2021
- Museum, Queensland. "Coming of the Light - Torres Strait Islands." Queensland Museum, Queensland Government, 2021, www.qm.qld.gov.au/Explore/Find+out+about/Aboriginal+and+Torres+Strait+Islander+Cultures/Gatherings/Coming+of+the+Light+Torres+Strait+Islands. Accessed 5 August 2021
- Noriega, R., et al. "Torres Strait Fisheries." Department of Agriculture, Water and the Environment, Australian Government, 2020, www.agriculture.gov.au/abares/research-topics/fisheries/fishery-status/torres-strait. Accessed 13 August 2021
- Parker, Robert, and Helen Milroy. *Mental Illness in Aboriginal and Torres Strait Islander Peoples*, The Telethon Kids Institute, 2013, www.telethonkids.org.au/globalassets/media/documents/aboriginal-health/working-together-second-edition/wt-part-2-chapt-7-final.pdf. Accessed 9 August 2021
- Parliament, Queensland. "Indigenous Suffrage Timeline." Indigenous Suffrage Timeline Queensland, Queensland Parliament, 2015, www.parliament.qld.gov.au/documents/explore/education/factsheets/Factsheet_10.5_IndigenousSuffrageTimeline.pdf. Accessed 3 August 2021
- Pendleton, Linwood, et al. "The Great Barrier REEF: Vulnerabilities and Solutions in the Face of Ocean Acidification." *Regional Studies in Marine Science*, Elsevier, 10 July 2019, www.sciencedirect.com/science/article/pii/S2352485518306959#b15. Accessed 17 August 2021
- Queensland Government, Queensland Health. *Traditional Aboriginal and Torres Strait Islander Foods*, Queensland Government, n.d., www.health.qld.gov.au/data/assets/pdf_file/0032/428288/jarjums-sect3-less3.pdf. Accessed 14 August
- Queensland Government. *Torres and Cape Hospital and Health Service*, Queensland Health, n.d., www.health.qld.gov.au/torres-cape. Accessed 4 August 2021
- Readfearn, Graham. *Great Barrier Reef's Third Mass Bleaching in Five Years the Most Widespread Yet*, Guardian News and Media, 6 Apr. 2020, www.theguardian.com/environment/2020/apr/07/great-barrier-reefs-third-mass-bleaching-in-five-years-the-most-widespread-ever. Accessed 9 August 2021
- Reaka-Kudla, M. L. (1997). "The global biodiversity of coral reefs: a comparison with rain forests," in *Biodiversity II: Understanding and Protecting Our Biological Resources*, eds M. L. Reaka-Kudla, D. E. Wilson, and E. O. Wilson (Washington, DC: Joseph Henry Press), 83–108. Accessed 3 August 2021

- Reef 2050 Long-Term Sustainability Plan—July 2018, Commonwealth of Australia 2018. Fabricius, K.E., Neill, C., Van Ooijen, E. et al. Progressive seawater acidification on the Great Barrier Reef continental shelf. *Sci Rep* 10, 18602 (2020). <https://doi.org/10.1038/s41598-020-75293-1>. Accessed 4 August 2021
- Richards, Zoe T, and Jon C Day. “Biodiversity of the Great Barrier Reef-how adequately is it protected?” *PeerJ* vol. 6 e4747. 8 May. 2018, doi:10.7717/peerj.4747. Accessed 12 August 2021
- Sakakibara, Chie. “‘Our Home Is Drowning’: Iñupiat Storytelling and Climate Change in Point Hope, Alaska.” *Geographical Review*, vol. 98, no. 4, 2008, pp. 456–475. JSTOR, www.jstor.org/stable/40377348. Accessed 12 Aug. 2021. Accessed 9 August 2021
- Sala, Enric, and Nancy Knowlton. “Global Marine Biodiversity Trends.” *Global Marine Biodiversity Trends, Annual Review of Environment and Resources*, 2006, www.annualreviews.org/doi/full/10.1146/annurev.energy.31.020105.100235. Accessed 5 August 2021
- Schaffelke, Britta, and Ken Anthony. Crown-of-Thorns Starfish Research Strategy, Australian Institute of Marine Science, 2015, www.aims.gov.au/sites/default/files/AIMS%20CoTS%20research%20strategy.pdf. Accessed 13 August 2021
- Shepherd, Carrington C J et al. “Socioeconomic disparities in the mental health of Indigenous children in Western Australia.” *BMC public health* vol. 12 756. 10 Sep. 2012, doi:10.1186/1471-2458-12-756. Accessed 8 August 2021
- Shnukal, A, and G Ramsay. “Navigating Boundaries: The Asian Diaspora in Torres Strait.” *An Overview of Torres Strait Islander-Asian Contact*, Australian National University, 2004, press-files.anu.edu.au/downloads/press/n3973/pdf/ch02.pdf. Accessed 9 August 2021
- Sjöholm P, Pakkala K, Davison B, et al. Socioeconomic status, remoteness and tracking of nutritional status from childhood to adulthood in an Australian Aboriginal Birth Cohort: the ABC study. *BMJ Open* 2020;10: e033631. doi: 10.1136/bmjopen-2019-033631. Accessed 11 Aug 2021
- Skinner K., Pratley E. (2016) Indigenous Peoples and Food. In: Thompson P., Kaplan D. (eds) *Encyclopedia of Food and Agricultural Ethics*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6167-4_587-1. Accessed 14 August 2021
- Sully, S., Burkepile, D.E., Donovan, M.K. et al. A global analysis of coral bleaching over the past two decades. *Nat Commun* 10, 1264 (2019). <https://doi.org/10.1038/s41467-019-09238-2>. Accessed 8 August 2021
- Trewin D, Madden R. The health and welfare of Australia’s Aboriginal and Torres Strait Islander peoples. 2005. Canberra: Australian Bureau of Statistics. Accessed 18 August 2021
- TSIRC, Torres Strait Island Regional Council. Culture, Protocols & Heritage, Torres Strait Island Regional Council, n.d., www.tsirc.qld.gov.au/our-communities/culture-protocols-heritage. Accessed 14 August 2021
- U.S. Department of Arts and Culture. “Honor Native Land: A Guide and Call to Acknowledgment.” U.S. Department of Arts and Culture, n.d., usdac.us/nativeland. Accessed 5 August 2021

- United States Government, Environmental Protection Agency. "Effects of Ocean and Coastal Acidification on Marine Life." EPA, Environmental Protection Agency, n.d., www.epa.gov/ocean-acidification/effects-ocean-and-coastal-acidification-marine-life. Accessed 11 August 2021
- USDA, United States Department of Agriculture. Climate Change and Indigenous Peoples: A Synthesis of Current Impacts and Experiences, United States Department of Agriculture, 2014, www.fs.fed.us/pnw/pubs/pnw_gtr944.pdf. Accessed 14 August 2021
- USGS, U.S. Geological Survey. "Climate Change, Coastal Tribes and Indigenous Communities." Climate Change, Coastal Tribes and Indigenous Communities, U.S. Geological Survey, 2015, www.usgs.gov/news/climate-change-coastal-tribes-and-indigenous-communities. Accessed 12 August 2021
- Westcott, D.A., Fletcher, C.S., Kroon, F.J. et al. Relative efficacy of three approaches to mitigate Crown-of-Thorns Starfish outbreaks on Australia's Great Barrier Reef. *Sci Rep*10, 12594 (2020). <https://doi.org/10.1038/s41598-020-69466-1>. Accessed 14 August 2021
- Whitmore, James, et al. "The State of the Great Barrier Reef: Experts Respond." The State of the Great Barrier Reef: Experts Respond, The Conversation, 9 Aug. 2021, theconversation.com/the-state-of-the-great-barrier-reef-experts-respond-30450. Accessed 8 August 2021
- World Wildlife Fund (WWF). 2015. "Australia's Great Barrier Reef Under Threat." <https://www.worldwildlife.org/stories/australia-s-great-barrier-reef-under-threat>. Accessed August 2, 2021.



Using Networks and Supervised Machine Learning to Analyze and Predict the Results of Team Sports Games

David Li

Author Background: *David Li grew up in the United States and currently attends William P. Clements High School in Sugar Land, Texas in the United States. His Pioneer research concentration was in the field of computer science and titled “Graph Analytics: Algorithms and Machine Learning.”*

Abstract

As graph analysis methods have advanced, people have started using networks to analyze sports. However, because this idea is still fairly new, few bettors and sports analysts have used it in their work. This paper intends to answer the question of whether graphs could be useful in predicting the results of sports games in the National Football League (NFL), English Premier League (EPL), National Hockey League (NHL), and National Basketball Association (NBA). I answer this question by creating a network where nodes represent teams and edges represent games. I then use classification models to predict which team will win each game, regression models to predict the score difference of each game, and various centrality metrics to rank the teams in each league. The results showed that graphs were generally useful for predicting and analyzing results in sports games. One classification model performed worse than randomly guessing, but the other models, particularly when used on NFL data, were very accurate. The results when predicting score difference were not as impressive, but still suggest that it would be beneficial for sports analysts to consider network structure in their work.

1. Introduction

Graphs are very flexible data structures that are typically used to represent relationships between objects. They can represent various types of relationships, such as collaboration between writers, interaction between users on social media sites, the flow of information across the web, borrowing between banks, and more [1].

As people have developed better algorithms to learn about and

manipulate graphs, the idea of using networks to analyze social structures has become increasingly popular, finding countless applications in countless fields. Engineers have used graphs to design circuits, sociologists have used graphs to measure the popularity of actors, biologists have used graphs to model protein interactions, and chemists have used graphs to study atomic structure.

To learn more about these graphs, scientists began using machine learning on graphs. This application of machine learning is used by search engines to recommend websites and by social media sites to recommend friends, among many other things. Applying machine learning techniques to graphs allows us to predict what edges will form in the future, classify nodes, classify edges, use graph neural networks, more [2].

In this paper, I will use network analysis and machine learning to analyze and predict the results of the National Football League (NFL), National Hockey League (NHL), National Basketball League (NBA), and English Premier League (EPL). I will analyze the structure of various sports graphs and perform various measurements to learn more about each sport. I will then use the graphs to examine how various features, such as the number of injuries in a sport, the average score per game of a sport, and other metrics, impact graph structure.

From there, I will predict the results of games in the NFL, NHL, NBA, and EPL using machine learning models based purely on data I obtained from the graphs. I will predict not only which team will win each game, but also by how much they will win the game. The problem of predicting who wins is similar to the problem of predicting edge direction in directed graphs or predicting edge signs in signed graphs, which can be treated as a classification problem [3]. The problem of predicting score difference is similar to the problem of predicting edge weights in graphs, which can be treated as a regression problem [4]. By predicting results using solely data from graphs, I aim to determine how useful graph data could be to the field of sports analysis, and whether it would be beneficial for sports analysts and bettors to consider sports graphs as they analyze games.

2. Related Work

In one paper, McLean et al. [5] used a graph where nodes represented soccer players and edges represented passes between them. They used this graph to measure degrees of connectivity between nodes, a clustering coefficient, betweenness centrality, and other metrics, in order to analyze the play style of each team. They then used this data to answer more questions, such as where certain players are the most effective, how they interact with their teammates, and what type of player they are (i.e., stars or supporting players). This data could then be used to help coaches decide where and how to play their players. In another paper, Arriaza-Ardiles et al. [6] used the same type of graph but measured different quantities. In this paper, they measured density, cohesion, and degree centrality, and used various data analysis methods to identify which parts of the pitch teams played the best in, the quality of relationships between various players on the pitch, and differences between how teams played in the group stages vs. the knockout stages of tournaments. In another paper, Korte and Lames [7] used similar graphs for multiple sports and calculated various centrality measurements. They then used those metrics, along with a minimum

spanning tree, to determine the most influential positions in various sports teams.

In all three of these papers, the authors used the same type of graph but calculated a variety of different metrics to analyze different aspects of a single sport, which shows how effective network analysis can be in the sports world. I will use networks in a similar way to find practical insights about sports. However, instead of using networks to dive deeply into one specific sport, I will use networks to analyze differences across multiple sports.

In another paper, Siranart et al. [8] used many rankings systems, a logistic regression model, and a random forest model, with networks, to predict sports results. They treated it as a classification problem, where the results were either win, draw, or loss. They ended up with fairly positive results, with the metrics predicting the results of games with much better accuracy than random guesses almost every time. However, they used a very limited number of models. I aim to use a much larger set of models to potentially identify the best ones and see if that trend is true for all models. Furthermore, Siranart et al. only predicted whether a team would win, draw, or lose, not considering how much a team wins by. In my paper, I predict score difference to better understand the extent to which graphs are useful in sports prediction.

3. Background

3.1. Definitions

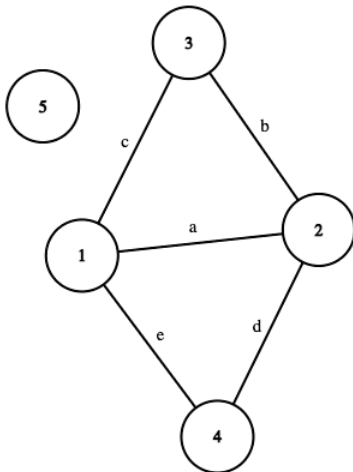


Figure 1: A picture of an undirected graph

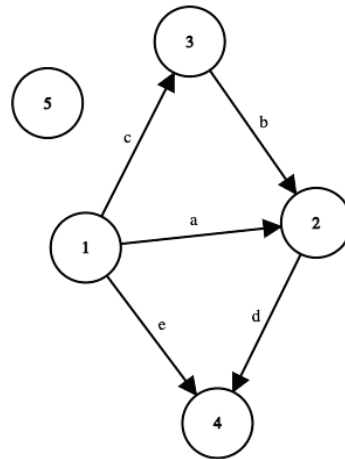


Figure 2: A picture of a directed graph

- **Graph** – A graph is a data structure that specifies the relationship between multiple objects. It consists of a set of objects called nodes. Certain pairs of these nodes are connected by links called edges.

Sometimes, these edges only point in one direction, and sometimes they point in both directions. If the edges in a graph only point in one direction, then the graph is considered a directed graph. Otherwise, it is an undirected graph. If the edges have weights on them, then the graph is considered a weighted graph. Otherwise, it is an unweighted graph [9]. In Figures 1 and 2, a, b, c, d, and e are edges, while 1, 2, 3, 4, and 5 are nodes.

- **Neighbor** – In a *graph*, the *neighbors* of a *node* are the *nodes* that are connected to a specified *node* by an *edge*. Lists of outgoing *neighbors* of a *node* u are often represented by $N^+(u)$ while lists of incoming *neighbors* are often represented by $N^-(u)$ [9]. In Figure 1, *node* 3 has two *neighbors*, namely *nodes* 1 and 2. In Figure 2, *node* 2 has two incoming *neighbors*, namely *nodes* 1 and 3. *Node* 2 only has one outgoing *neighbor*, namely *node* 4.
- **Degree** – The *degree* of a *node* is the number of *edges* connected to it. In a *directed graph*, the *in-degree* is the number of *edges* pointing towards the *node*, and the *out-degree* is the number of *edges* pointing away from the *node*. In a *weighted graph*, the *strength* of a *node* is the sum of the *weights* of the *edges* connected to a *node*. *Inward* and *outward strength* are defined similarly to *in-degree* and *out-degree*, respectively [9].

$$\begin{aligned} str_{out}(u) &= \sum_{v \in N^+(u)} w_{u,v} \\ str_{in}(u) &= \sum_{v \in N^-(u)} w_{v,u} \end{aligned}$$

Here *str* represents the *strength* of a *node*, $w_{i,j}$ is the *weight* of an *edge* from node i to node j , $N^+(u)$ is the list containing the *outgoing neighbors* of node u , and $N^-(u)$ is the list containing the *incoming neighbors* of node u .

3.2. Metrics and Ranking Systems

- **Weighted PageRank** – The *PageRank* algorithm was originally developed by Google to rank websites but can also be used to rank nodes in graphs [10]. In this algorithm, every node is given an initial *PageRank* value, or $PR(i)$, of $1/n$, where n is the number of nodes in the graph. We then run k iterations of an algorithm, where k is a runtime parameter. In each iteration, each page divides its *PageRank* value equally throughout its neighbors. Each page then updates its value to the sum of shares that it received. Scores are then scaled down by a factor of s and the residual $1 - s$ units of *PageRank* are then redistributed equally among all nodes.

We can apply this to weighted graphs by modifying the original *PageRank* algorithm slightly. Instead of distributing a page's *PageRank* value equally among its neighbors, we can distribute *PageRank* to its

neighbors proportionally to the weights of the edges connecting it to its neighbors. We end up with the following function [11]:

$$WPR(u) = \frac{1-s}{n} + s * \sum_{v \in N^-(u)} (WPR(v) * \frac{w_{v,u}}{str_{out}(v)})$$

Here WPR is the weighted *PageRank* function, n is the number of nodes in the graph, s is the scaling factor, $w_{v,u}$ is the weight of the edge from v to u , $N^-(u)$ is the list containing the incoming neighbors of node u , and str_{out} is the outwards strength of a node.

- **VoteRank** – *VoteRank* is an algorithm suggested by Zhang et al. [12] that can be used to rank nodes in a graph. Initially, every node is given a pair of values – s_u and va_u . s_u is the voting score of a node, which denotes the number of votes node u has received from its neighbors. va_u is the voting ability of a node and denotes the number of votes that node u can give its neighbors. Initially, s_u and va_u are set to 0 and 1, respectively. Then, the following series of steps occurs:

1. All nodes vote for their neighbors, and the voting score of each node is recalculated.
2. The node with the highest voting score v_{max} has its voting ability set to 0. This prevents it from voting in any future turns.
3. We weaken the voting abilities of the nodes that voted for v_{max} .
4. We repeat steps 1 through 3 until r nodes have been selected, where r is a runtime parameter.

The order in which nodes are selected is the order by which the nodes are ranked.

- **Clustering Coefficient** — The clustering coefficient of a node is a measure of how clustered the nodes in a graph are around a specified node. There are several ways to represent the clustering coefficient in a weighted directed graph [13]. However, I will use an unweighted clustering coefficient in this paper to better represent the graph structure. In this paper, I will use the following definition [14]:

$$c_u = \frac{2T(u)}{\deg(u) (\deg(u) - 1)}$$

Here, c_u is the clustering coefficient of node u , $T(u)$ is the number of triangles through node u , and $deg(u)$ is the degree of node u .

We can also find the clustering coefficient of a graph with the following equation:

$$C = \frac{1}{n} \sum_{v \in G} c_v$$

Here, C is the average clustering coefficient of the graph, n is the number of nodes in the graph, and c_v is the clustering coefficient of node v .

- **Accuracy** – I will use two definitions of accuracy in this paper.

One definition of accuracy I will use is the following:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Another definition of accuracy I will use is the following:

$$\text{Accuracy} = \frac{\sum_{\text{all predictions}} \text{predicted value} - \text{actual value}}{\text{Total Number of Predictions}}$$

3.3. Machine Learning Methods

- **Linear Regression** — The linear regression model models the relationship between a group of independent variables and one dependent variable using linear prediction functions. It assumes that the relationships between the independent variables and the dependent variable are linear. It is typically used for modeling regressions but can also be modified and used for classification purposes [15].
- **Logistic Regression** — The logistic regression model models the relationship between a group of independent variables and one binary dependent variable using a logistic function. It is typically used for binary classification [15].
- **Support Vector Machine (SVM)** — The SVM model assigns data to categories in space so as to maximize the distance between the categories. It can be used to make both classification and regression models [15].
- **Quadratic Discriminant Analysis (QDA)** — The QDA classifier classifies vectors by mapping them in open space. It then uses a quadratic shape in space derived from the training data to classify the vector [15].
- **Naive Bayes** — The naive Bayes classifier is a probabilistic classifier based on applying Bayes' theorem while assuming that all the independent variables are very independent of each other [15].
- **Decision Tree** — The decision tree model uses a tree-like model of various decisions to map an input vector to some class or value. It can be used for classification or regression [15].

- **Random Forest** — The random forest model uses various random decision tree models and averages their results. As a result, they generally outperform decision trees and can also be used both for classification and regression [15].
- **K-Nearest Neighbors (KNN)** — The KNN algorithm can be used for both classification and regression purposes. It groups objects based on the distance between them when plotted in space [15].
- **Gaussian Process** — The Gaussian Process regression model calculates the probability distribution over all possible functions that fit the data instead of directly finding a parametric function [15].
- **Multi-Layer Perceptron (MLP)** — The MLP is a type of Artificial Neural Network. The model can be used for both regression and classification [15].
- **Adaptive Boosting (ADA)** — The adaptive boosting model is a boosted classifier that uses the results of various other models to make a prediction [15].

4. Technical Approach

4.1. Network Structure

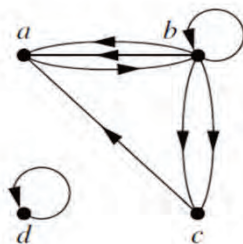


Figure 3: A picture of a directed multigraph

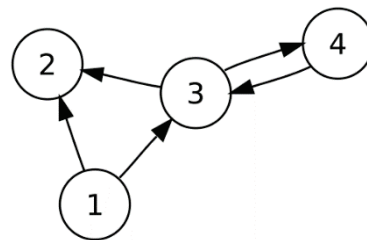


Figure 4: A picture of a simple directed graph

In this paper, I will use two main types of graphs, alternating depending on what is best for each algorithm. For both networks, nodes are teams' names, and edges from node a to node b represent team a winning against team b .

The first type of graph I use is a *directed multigraph*. In this graph, every game g is represented by an edge directed from the team that wins to the team that loses. Edges are weighted with the following function:

$$W_g = \frac{s_{g_i} + 1}{s_{g_j} + 1}$$

Here, g represents a game, s_{g_u} is the number of points that team u scored in game g , team i is the team that won, and team j is the team that lost.

The general idea here is that the weight of an edge is the ratio between the scores of the two teams. This is a good way to weight the graph because it provides a good idea of how much a team wins by. Another possible way to weight the graph would be by score difference, but that would be misleading because a 100–90 victory and a 10–0 victory should not be treated the same. The 10–0 victory should be seen as more significant. We add 1 to each score to make sure that we do not divide by 0.

The other type of graph I use is a *simple directed graph*. In this graph, every pair of nodes i and j has two edges between them. One extends from i to j , and the other from j to i . An edge from any node i to any node j stores data about all the games where team i won against team j . In this graph, edges are weighted by the following function:

$$W_{ij} = \sum_{g \in G_{ij}} W_g$$

Here, W_{ij} represents the weight of the edge from i to j , g represents a game, G_{ij} is the set of games where team i won against team j , and W_g represents the weight that game g would have on the multigraph.

In the *simple directed graph*, edges from any node i to any node j will also store the number of games where team i won against team j .

In this paper, I use the *directed multigraph* to count triangles and edges, and the *simple directed graph* to get VoteRanks, the weighted out-degree, and the clustering coefficient. I will use a reversed copy of the *simple directed graph* to get PageRank values because nodes with many incoming edges receive a high PageRank and wins in our original graph are outgoing instead of incoming.

4.2. Datasets

I will be using a series of datasets from Kaggle.com, an online database of user-submitted datasets. Details about the datasets I will use are listed below.

- **NBA Games Data [16]** — This dataset contains detailed information on every NBA game from the 2003/2004 to the 2020/2021 season. It includes the dates of the games, which team played at home, which team played away, how many points each team scored, how many rebounds each team got, and more. I will use data from the 2011/2012 season to the 2015/2016 season.
- **Premier League [17]** — This dataset contains basic information about every EPL game from the 2006/2007 season to the 2017/2018 season. It includes the results of each game, the team that played at home, the team

that played away, and the number of goals each team scored. I will use data from the 2011/2012 season to the 2015/2016 season.

- **NHL Games Data [18]** — This dataset includes extremely detailed information about every NHL game from the 2000/2001 season to the 2019/2020 season. It contains every metric that the NFL measures, including the time of the game, the team that played at home, the team that played away, how many points the teams scored, and more. I will use data from the 2011/2012 season to the 2015/2016 season.
- **NFL Scores and Betting Data [19]** — This dataset includes information about every game from the 1978/1979 season to the 2020/2021 season. It includes the results of each game, the team that played at home, the team that played away, the number of points each team scored, and more. I will use data from the 2011/2012 season to the 2015/2016 season.

Each dataset is formatted slightly differently and contains a significant amount of extra data, so I preprocessed it, leaving only the teams and the scores.

In my paper, I combine data from the 2013/2014 season and 2014/2015 season in one dataset, which I often refer to as the “smaller dataset.” I also combine data from the 2011/2012 season up to the 2015/2016 season in another dataset, which I often refer to as the “larger dataset.”

4.3. Prediction Models

- **Ranking-Based Models**
 - **Baseline** — The Baseline model will randomly guess the results with equal probability. The EPL and NFL have draws, while the NBA and NHL do not. However, there have only been nine draws in the NFL in the last 12 years, so I have chosen for my baseline model to ignore the draws. For the EPL, the model will guess win, loss, or draw. For the NFL, NBA, and NHL, the model will guess either win or loss.
 - **PageRank** — In this model, I will look at the PageRank values of the two teams in question and give the victory to the team with the higher PageRank value. If they have the same PageRank value, I will output a draw. To calculate the PageRank values of the teams, I will use a reversed copy of the simple directed graph.
 - **Out-Degree** — In this model, I will look at the weighted out-degrees of the two teams in question and give the victory to the team with the higher value. If they have the same out-degree value, I will output a draw. I will compute this value using the simple directed graph.
 - **VoteRank** — In this model, we will look at the VoteRanks of the

two teams in question and give the victory to the team selected earlier. If neither has a VoteRank value, I will output a draw. I will compute this value using the simple directed graph.

- **Edge Direction Prediction Models**

The features I will give the edge direction prediction models include the PageRank values of both teams, the weighted out-degrees of both teams, the VoteRanks of both teams, the number of paths of length one between the teams, and the number of paths of length two between the teams. I give the models the various centrality metrics/rankings to help the model predict which team is generally better. The paths of length one between the two teams represent the direct record between the two teams, which helps the model identify any interesting relationships between the two teams. We give paths of length two between two nodes because a path of length two from node a to node c would suggest that team a should defeat team c . For example, if there were a directed edge from node a to node b , as well as one from node b to node c , that would suggest that there should be an edge from node a to node c . I tried to add the number of paths of lengths three and four, but the program took too long to run.

The models I will test are the following: Logistic Regression, Naive Bayes, Support Vector Machine, K-Nearest Neighbors, Random Forest, Gaussian Process, Adaptive Boosting, Multi-Layer Perceptron, and Quadratic Discriminant Analysis.

- **Edge Weight Prediction Models**

In the edge weight prediction models, I will use the same features I used for the edge direction prediction task. This time, however, I will not predict who wins or loses. Instead, I will predict the score difference in games, which is equivalent to predicting the weight of an edge. I will predict the weight of a directed edge going both ways. I will use the average difference between my predicted values and the actual values to measure accuracy.

The models I will test are the following: Linear Regression, Decision Tree, Support Vector Machine, Gaussian Process, Multi-Layer Perceptron, K-Nearest Neighbors, Adaptive Boosting, and Random Forest.

5. Experiments

5.1. Experimental Setup

I ran the experiments below using Python 3.8.2 on a 2017 MacBook Air with 16GB of memory, a 1.4 GHz Dual-Core Intel Core i7 processor, and macOS Big Sur version 11.5 installed. I ran my code using Visual Studio Code. I used the NetworkX package to create graphs and run graph algorithms, the pandas package to model data, the scikit-learn package to create classification and

regression models, and the Matplotlib package to create scatterplots to visualize the data.

5.2. Graph Structure

Table 1: *Data for the 2014/2015 season*

Metric	EPL	NFL	NBA	NHL
Nodes (Teams)	20	32	30	30
Edges (Games)	380	267	1418	1319
Clustering Coefficient	0.3977	0.3243	0.8116	0.8162
Triangles	695	177	26588	27148
Average Points Per Game	2.47	46.97	196.10	6.84
Maximum Players on Roster	25	53	15	23

As we can see from Table 1, which contains basic data about the multigraph representations of the 2014/2015 season for each league, as well as basic information for each league, the NBA and the NHL both have significantly more games per season than do the EPL or NFL. It is crucial that we take this into account as we analyze our data.

The number of triangles in a graph represents the number of combinations of games such that team a beat team b , team b beat team c , and team c beat team a . These patterns are a sign of inconsistency.

We see that, even though the EPL only has about 1.3 times the number of games as the NFL, the EPL has almost quadruple the number of triangles. When we consider that the EPL also has draws, meaning that many of the EPL games cannot contribute to any triangles, this difference is even more dramatic. This makes sense given that an NFL team's roster consists of 53 players, while an EPL roster contains no more than 25 players. This means that in the EPL, each individual player is more important to their team's result. If one EPL player gets injured or has a bad game, it impacts their team's performance much more than if one NFL player gets injured or has a bad game. This suggests that EPL results should be less consistent than NFL games. Furthermore, the EPL has far fewer points per game than the NFL. This means that one lucky play can potentially give a team a huge advantage. This makes it more likely for the "worse" team to pull an upset and win a game. This suggests that the prediction models should perform better on the NFL than on the EPL.

We also see that the NBA has slightly more games but slightly fewer triangles than the NHL. This suggests that the NHL is slightly less consistent than the NBA. This makes sense in the real world as well. Although an NBA roster consists of no more than 15 players while an NHL team has 23 players, an NHL game has an average of fewer than seven goals per game, while NBA games average nearly 200 points per game. These two factors in combination mean that the two leagues should have similarly inconsistent results, which is illustrated in the data. This suggests that the prediction models should perform similarly on the NBA and NHL data.

Another interesting factor is the clustering coefficients of the graphs. The clustering coefficients tend to follow the same trend as the triangles, with the EPL having a significantly higher clustering coefficient than the NFL, and the NBA

having a slightly higher clustering coefficient than the NHL. This also makes sense, as the more inconsistent a league is, the more clustered the edges should be in the directed multigraph.

5.3. Ranking Systems

Table 2: Top rankings predicted by the ranking models and actual rankings for the 2014/2015 season. The numbers in parentheses represent the actual rankings of the teams that season. The NFL ranks teams by win percentage with some tiebreakers. However, teams across conferences are not ranked against each other, and multiple teams can have the same win percentage, so I gave teams with the same win percentage the same rank.

PageRank	VoteRank	Out-Degree	Actual
NHL			
Lightning (5)	Rangers (1)	Rangers (1)	Rangers (1)
Blackhawks (7)	Red Wings (10)	Lightning (5)	Canadiens (2)
Rangers (1)	Blackhawks (7)	Blackhawks (7)	Ducks (3)
NFL			
Patriots (1)	Patriots (1)	Rams (20)	Patriots (1)
Seahawks (1)	Cowboys (1)	Colts (6)	Cowboys (1)
Chiefs (14)	Seahawks (1)	Dolphins (15)	Seahawks (1)
EPL			
Chelsea (1)	Chelsea (1)	Man. City (2)	Chelsea (1)
Man. City (2)	Man. United (4)	Chelsea (1)	Man. City (2)
West Brom. (12)	Arsenal (3)	Arsenal (3)	Arsenal (3)
NBA			
Warriors (1)	Warriors (1)	Warriors (1)	Warriors (1)
Cavaliers (4)	Hawks (2)	Hawks (2)	Hawks (2)
Rockets (3)	Rockets (3)	Cavaliers (4)	Rockets (3)

Next, we use the simple directed graph representations of the leagues to predict the rankings of the leagues.

Logically, it would follow that the ranking systems (PageRank, VoteRank, Out-Degree) would rank the teams with better records higher. However, as we can see from Table 2, which shows the top predicted seeds using the three ranking systems using data from the 2014/2015 season, this is not always the case.

One main reason is that a 1–0 victory is valued similarly to a 10–0 victory in real life. However, in our model, a 10–0 win is markedly more significant than a 1–0 win. This allows for bad teams with lucky games to rank very highly using the graph rankings systems. This is precisely why in Table 2, West Bromwich Albion (West Brom) is predicted to be ranked so highly in the PageRank rankings for the EPL in the 2014/2015 season. In that season, they had a 4–0 victory and multiple 3–0 victories, with one of the 3–0 victories being over Chelsea, who won the league that season.

Another reason the predictions are off is that the leagues all have different ranking systems and do not play the same number of games against every other team. For example, in the EPL, teams play twice against every other

team—once at home and once away. They are then ranked by points, where a win gives 3 points, a draw gives 1 point, and a loss gives 0 points. This is dramatically different from the NBA, where teams play teams in the other conference twice, teams from their own division four times, and teams from their conference but not their division three or four times. They are then ranked by win percentage. In the NHL, teams are ranked by points and play teams from their own division three or four times, teams from their conference but not division three times, and the teams in the other conference twice. In the NFL, teams are ranked by win percentage and teams play the three other teams in their division twice, four teams from one American Football Conference (AFC) division once, four teams from one National Football Conference (NFC) division once, two intraconference games, and one interconference game. These differences are likely a reason for the difference in the results.

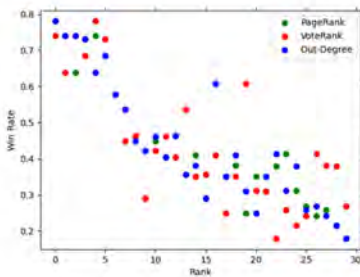


Figure 5: *EPL Rankings*

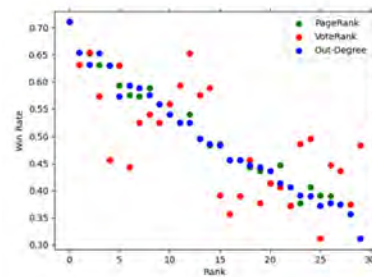


Figure 6: *NBA Rankings*

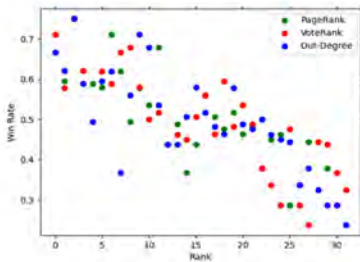


Figure 7: *NFL Rankings*

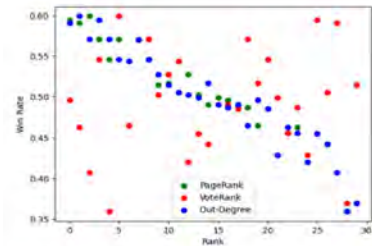


Figure 8: *NHL Rankings*

Figures 5, 6, 7, 8: *Ranks of teams as predicted by ranking systems with their winning percentages*

The Out-Degree model seems to perform the worst when it comes to ranking the top teams. This is not surprising, as it only considers the wins that a team receives, ignoring their losses. However, when we look at more data points, as we see in Figures 5-8, which contain scatterplots with the full ranking predictions, the Out-Degree models appear quite accurate.

If we look at all the data, as we see in Figures 5-8, we see that VoteRank actually has the worst overall prediction by far. In the NHL ranking prediction graph, the VoteRank rankings almost appear to be completely random. However, if we only look at the predictions for the three highest ranked teams, VoteRank seems to be one of the best, predicting both the NFL's and NBA's top

three seeds perfectly. This suggests that VoteRank performs well at identifying the most significant nodes in a graph but struggles when it comes to ranking the remaining nodes. This also suggests that VoteRank will not be very good at predicting the results of games.

If we look at the PageRank and Out-Degree models, the graphs seem to suggest that the NBA's seeding system is the most accurate, with the models nearly generating straight lines. The ranking systems seem to do the worst when it comes to NFL prediction, suggesting that the NFL seeding system tends to overvalue certain teams and undervalue others. This is not surprising, considering the NFL's extreme division-based structure, which gives teams in bad divisions significant advantages and teams in good divisions significant disadvantages. If every team in the leagues played the other teams with the same frequency, then we would likely get better results.

Overall, the rankings algorithms seemed to perform well when predicting rankings, with PageRank and Out-Degree being the best metrics overall.

5.4. Edge Direction Prediction

Table 3: *Larger Dataset: Results for the 2011/2012 season to the 2015/2016 season. Bold numbers represent the best results for each league, while italicized numbers represent the worst non-baseline results for each league. A red number represents a result worse than the baseline.*

Model	EPL	NFL	NBA	NHL
Baseline Model	33.3	50.0	50.0	50.0
PageRank Model	51.2	63.4	62.5	56.7
VoteRank Model	49.6	61.9	58.3	<i>50.0</i>
Out-Degree Model	51.5	62.1	62.3	56.6
Logistic Regression Model	53.3	65.6	63.9	58.3
Naïve Bayes Model	51.2	61.4	61.9	57.9
Support Vector Machine Model	52.6	58.4	63.5	55.3
Quadratic Discriminant Analysis Model	50.1	65.3	63.1	57.7
K-Nearest Neighbors Model	49.1	62.9	60.0	54.3
Random Forest Model	46.7	61.3	60.6	53.4
Gaussian Process Model	42.5	58.4	60.0	57.9
Adaptive Boosting Model	52.5	47.9	63.5	57.9
Multi-Layer Perceptron Model	<i>40.4</i>	62.6	63.1	54.5
Decision Tree Model	41.1	50.1	<i>57.2</i>	52.6

Table 4: *Smaller Dataset: Results for the 2013/2014 season to the 2014/2015 season. Bold numbers represent the best results for each league, while italicized numbers represent the worst non-baseline results for each league.*

Model	EPL	NFL	NBA	NHL
Baseline Model	33.3	50.0	50.0	50.0
PageRank Model	50.8	68.5	64.6	58.6
VoteRank Model	49.7	65.5	61.3	55.9
Out-Degree Model	49.6	66.3	65.0	58.4
Logistic Regression Model	53.3	72.4	67.2	58.1
Naïve Bayes Model	50.1	74.6	65.7	58.0
Support Vector Machine Model	50.0	<i>56.7</i>	<i>55.6</i>	59.5
Quadratic Discriminant Analysis Model	49.5	71.2	65.3	59.2
K-Nearest Neighbors Model	45.8	69.4	60.0	<i>54.3</i>
Random Forest Model	46.3	61.9	60.8	56.1
Gaussian Process Model	42.1	61.2	57.0	54.6
Adaptive Boosting Model	53.2	58.2	65.5	59.0
Multi-Layer Perceptron Model	50.0	73.9	59.4	58.5
Decision Tree Model	<i>38.4</i>	65.6	57.2	55.5

As we can see in Tables 3 and 4, which show edge direction prediction accuracies for the larger and smaller datasets, respectively, the models performed the worst on the EPL, which is expected, as the EPL allows for wins, draws, and losses, while the NBA, NFL, and NHL only have wins and losses. It is important to keep this in mind when comparing its results to the results of the other leagues in order to avoid false over-comparison.

Out of the other three leagues, the models tended to perform the worst on the NHL. This makes sense, as the NHL has a very low number of points per game and players per team, and thus has less consistent results, as illustrated earlier.

Previously, when we analyzed the graph structure, we were unable to compare the consistency between the NFL and the NBA due to the massive difference in games played. The new data, which shows the NFL models performing better than the NBA models in almost every situation, suggests that NFL games are slightly more consistent than NBA games. Surprisingly, even though the NFL seemed to have the most accurate results overall, it was the only league with a model that performed worse than the baseline.

One interesting data point comes in the EPL results data. For the NFL, NBA, and NHL, most of the models performed slightly better on the smaller dataset, which was expected because the smaller dataset only used two seasons of data, while the large dataset used five seasons of data, giving teams lots of time to improve or deteriorate. However, many of the EPL models performed better on the larger dataset than on the smaller dataset. At first glance, this might seem strange, but in context, it makes sense. In the EPL, teams that historically perform well tend to continue to perform well. For example, Manchester United, Chelsea, and Arsenal have finished in the top five of the EPL for almost all of the last 10 seasons. In other leagues, such as the NBA, this is not the case. For example, the Phoenix Suns, who were in the 2020/2021 NBA finals, were the second-worst team in the NBA in the 2018/2019 season. This

means that the results from two seasons of the EPL are more similar to EPL results over five seasons than results from the NBA, NFL, or NHL would be. With this in mind, it makes sense that the EPL models performed better on the larger dataset, because there is more data from which to learn.

Overall, the logistic regression model tended to perform the best, with it being the best model for six of the eight datasets. There did not seem to be a definitive worst model, with the decision tree model, support vector machine model, k-nearest neighbors model, multi-layer perceptron model, adaptive boosting model, and VoteRank model all being the worst model for at least one of the datasets. Out of the centrality metrics, the VoteRank model tended to perform the worst, with it being the least accurate ranking system-based model for four of the eight datasets.

One surprising observation was the performance of the adaptive boosting model. Logically, the models should perform the best on the NFL, which, we have demonstrated, shows the most consistent results. However, the ADA model performed worse on the NFL datasets than on both the NHL and NBA datasets.

Another surprising observation was the performance of the support vector machine model. Except for the EPL, the models should perform better on the smaller dataset than the larger dataset. However, the support vector machine model performed worse on the small dataset than the large dataset for the NFL, EPL, and NBA. This suggests that the SVM model performs well on large datasets but struggles on small datasets. This would explain the SVM's relatively NFL's poor prediction rates, as the NFL has much fewer games than the NBA and NHL. It would also explain why the SVM model did worse on the smaller NFL, NBA, and EPL datasets than on the larger ones. The Gaussian process model also seemed to perform better on the larger dataset, likely for similar reasons.

The decision tree model also yielded interesting results. For the EPL, NBA, and NHL, the model gave similar numbers on both the small and large datasets. However, the model was a shocking and substantial 15% more accurate on the smaller dataset than the larger dataset for the NFL data.

As we can see from the data, the prediction models are almost always better than the baseline model, which is expected. The data, especially some of the NFL data, was very impressive. However, some of the other models are only about 5% better than the baseline model, with one model even performing worse than the baseline model, which is slightly disappointing. Overall, the data suggest that graphs are useful for predicting results and that they should be considered by sports analysts.

5.5. Edge Weight Prediction

Table 5: Larger Dataset: Average error for the 2011/2012 season to the 2015/2016 season. Bold numbers represent the best results for each league, while italicized numbers represent the worst results for each league.

Model	EPL	NFL	NBA	NHL
Decision Tree Model	1.82	15.3	13.1	2.5
Decision Tree Model (Normalized)	1.33	1.29	1.18	1.25
Random Forest Model	1.29	11.6	10.7	2.04
Random Forest Model (Normalized)	0.95	0.97	0.97	1.02
Linear Regression Model	1.19	10.2	9.94	1.88
Linear Regression Model (Normalized)	0.87	0.86	0.90	0.94
Support Vector Machine Model	1.21	11.1	10.1	1.85
Support Vector Machine Model (Normalized)	0.88	0.91	0.90	0.92
Gaussian Process Model	1.55	12.6	12.8	2.32
Gaussian Process Model (Normalized)	1.14	1.07	1.16	1.16
Multi-Layer Perceptron Model	1.28	11.0	9.96	1.92
Multi-Layer Perceptron Model (Normalized)	0.98	0.97	0.91	0.98
K-Nearest Neighbors Model	1.30	11.36	10.8	1.96
K-Nearest Neighbors Model (Normalized)	0.95	0.96	0.98	0.98
Adaptive Boosting Model	1.24	10.7	9.99	1.89
Adaptive Boosting Model (Normalized)	0.91	0.90	0.90	0.95

Table 6: Smaller Dataset: Average error for the 2013/2014 season to the 2014/2015 season. Bold numbers represent the best results for each league, while italicized numbers represent the worst results for each league.

Model	EPL	NFL	NBA	NHL
Decision Tree Model	1.85	13.9	12.9	2.42
Decision Tree Model (Normalized)	1.32	1.15	1.16	1.23
Random Forest Model	1.37	10.4	10.6	2.00
Random Forest Model (Normalized)	0.98	0.85	0.96	1.00
Linear Regression Model	1.24	9.68	9.59	1.86
Linear Regression Model (Normalized)	0.88	0.81	0.87	0.93
Support Vector Machine Model	1.41	11.4	10.6	1.85
Support Vector Machine Model (Normalized)	0.91	0.94	0.90	0.93
Gaussian Process Model	1.51	12.1	12.0	2.19
Gaussian Process Model (Normalized)	1.07	1.01	1.09	1.10
Multi-Layer Perceptron Model	1.32	9.70	9.68	1.87
Multi-Layer Perceptron Model (Normalized)	0.95	0.84	0.89	0.93
K-Nearest Neighbors Model	1.33	10.9	10.6	2.04
K-Nearest Neighbors Model (Normalized)	0.95	0.91	0.97	1.02
Adaptive Boosting Model	1.29	9.74	9.76	1.88
Adaptive Boosting Model (Normalized)	0.93	0.83	0.89	0.95

In Tables 5 and 6, which show average error in score difference prediction on the larger and smaller datasets, respectively, it is difficult to compare the direct differences between the prediction models and reality because the sports have vastly different amounts of points per game. To account for this, I normalized

all the results by dividing the scores by the average score difference for their respective sports over the time period in question.

Looking at the results, it seems like the models performed similarly on all the leagues. The normalized error was ranged from 0.81 to 1.33, which is quite disappointing. This means that on average, the regression models predict a score difference that is almost an entire score difference away from the actual score difference. In other words, if a team won by 10 points in real life, the models might predict that the team won by 20 points or drew, which is quite poor. However, it is understandable that the predictions are not very good, as there are many factors, such as random lucky plays, off days, and more, that are unpredictable and are liable to impact the scores.

Overall, the models seemed to perform slightly better on the smaller dataset. This makes sense because the smaller dataset contains data over a shorter span of time, giving teams less time to get better or worse.

However, this is not the case for the EPL datasets, where three of the four models performed better on the larger dataset than the smaller dataset. We saw the same thing in the results of the edge direction prediction models, and it likely occurred for a similar reason—the EPL results get more consistent over time.

The support vector machine model also went against the trend of having better results for the smaller dataset, having either the same or better results on the larger dataset. This suggests that the support vector machine works better on larger datasets. While the smaller dataset was consistent, it likely did not have enough information to generate a good model.

Out of the models tested, the decision tree model performed the worst by far. It was the least accurate model for every league on both datasets. The linear regression model was the best, beating the other models in both datasets for the EPL, NFL, and NBA, while the support vector machine model did the best on the NHL datasets.

6. Conclusion

By using graphs to model the results of various sports leagues, I came to many interesting conclusions. The structure of the graphs told us that the EPL tended to have less consistent results than the NFL, and that the NHL and NBA had similarly consistent results. We validated these trends with real-life data, demonstrating that graphs are a viable way to model sports. When we used purely graph information to predict who would win or lose a game, we got relatively accurate results, especially with the NFL. However, our results with predicting the score difference were not as impressive, which is expected because it is a more challenging task. Still, the fact that we were able to achieve decent results while only using data from the graphs shows that graphs are potentially a very useful tool in the field of sports analysis. If sports bettors and analysts combined the data we got from the graphs with the other data from real life that they usually use, such as the players in the starting lineup, the styles of the two teams, the coaches, and more, then they could potentially generate much better data. Similarly, if we added non-graph data to our models, we could generate more accurate models.

The linear regression and logistic regression models performed the best

on the datasets used in this experiment. The ranking predictions suggested that VoteRank was a less reliable method of ranking nodes than PageRank and Out-Degree.

One possible way I could have improved the results of this experiment is by using data from college sports. I used data ranging from about 500 to 7000 games, which is a fair amount. However, this data ranged over 5 seasons, which gave teams a lot of room to improve or worsen. Furthermore, there were only 20–30 teams in each dataset. College sports would offer a lot more data overall, which would help generate better machine learning models. Furthermore, because there are so many college games every year, I could conceivably use data from a single season. This would make the data more consistent overall and give me much more accurate predictions.

I could also improve the edge weight prediction model by using models that are designed for edge weight prediction, such as the Weighted Stochastic Block Model, as proposed by Aicher et al. [20] or Model R, as proposed by Hou et al. [21], rather than simple regression models.

References

1. D. Easley and J. Kleinberg, *Networks, Crowds, and Markets*. Cambridge University Press, July 2010.
2. W. L. Hamilton, R. Ying, and J. Leskovec, “Representation learning on graphs: Methods and applications,” CoRR, vol. abs/1709.05584, 2017.
3. S. Kumar, F. Spezzano, V. S. Subrahmanian, and C. Faloutsos, “Edge weight prediction in weighted signed networks,” in IEEE 16th International Conference on Data Mining, ICDM 2016, December 12-15, 2016, Barcelona, Spain. IEEE, 2016, pp. 221–230.
4. C. Fu, M. Zhao, L. Fan, X. Chen, J. Chen, Z. Wu, Y. Xia, and Q. Xuan, “Link weight prediction using supervised learning methods and its application to yelp layered network,” IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 8, pp. 1507–1518, 2018.
5. S. Mclean, P. Salmon, A. Gorman, M. Naughton, and C. Solomon, “Do intercontinental playing styles exist? using social network analysis to compare goals from the 2016 euro and copa football tournaments knock-out stages,” Theoretical Issues in Ergonomics Science, vol. 18, pp. 1–14, 04 2017.
6. E. Arriaza-Ardiles, J. Martín-González, M. Zuniga, J. Sánchez-Flores, Y. de Saa, and J. García-Manso, “Applying graphs and complex networks to football metric interpretation,” Human Movement Science, vol. 57, pp. 236–243, 2018.
7. F. Korte and M. Lames, “Characterizing different team sports using network analysis,” Current Issues in Sport Science, 04 2018.
8. N. Siranart, K. Truong, and V. Berges, “Player rankings and outcome prediction in tournaments,” pp. 1–10, 2016.
9. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. The MIT Press, 2009.
10. S. Brin and L. Page, “The anatomy of a large-scale hypertextual web search engine,” Computer Networks, vol. 30, pp. 107–117, 1998.

11. W. Xing and A. Ghorbani, "Weighted pagerank algorithm," Proceedings. Second Annual Conference on Communication Networks and Services Research, 2004., pp. 305–314, 2004.
12. J. Zhang, D. Chen, Q. Dong, and Z. Zhao, "Identifying a set of influential spreaders in complex networks," CoRR, vol. abs/1602.00070, 2016.
13. J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, and J. Kertész, "Generalizations of the clustering coefficient to weighted complex networks," Physical Review E, vol. 75, no. 2, Feb 2007.
14. D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" dynamics," Nature, vol. 393, pp. 440–442, 1998.
15. scikit-learn. [Online]. Available: <https://scikit-learn.org/stable/>
16. N. Lauga. (2021) NBA games data. Kaggle. [Online]. Available: <https://www.kaggle.com/nathanlauga/nba-games>
17. Z. Nalla. (2018) Premier league. Kaggle. [Online]. Available: <https://www.kaggle.com/zaeemnalla/premier-league>
18. M. Ellis. (2020) NHL game data. Kaggle. [Online]. Available: <https://www.kaggle.com/martinellis/nhl-game-data>
19. Spreadspoke. (2021) NFL scores and betting data. Kaggle. [Online]. Available: <https://www.kaggle.com/tobycrabtree/nfl-scores-and-betting-data>
20. C. Aicher, A. Z. Jacobs, and A. Clauset, "Learning latent block structure in weighted networks," Journal of Complex Networks, vol. 3, no. 2, p. 221–248, Jun 2014.
21. Y. Hou and L. B. Holder, "Deep learning approach to link weight prediction," in 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 1855–1862.



Imagining a Carbon-Neutral Future Through Standard Oil's Last Four Sisters Standing: Beyond Petroleum, Chevron, ExxonMobil, and Shell

Dorothy Chan

Author Background: *Dorothy Chan grew up in the United States and currently attends Brooklyn Technical High School in Brooklyn, New York in the United States. Her Pioneer research concentration was in the field of environmental studies and titled “Waste or Not Waste, That’s the Question—Solving Major Environmental Problems.”*

Abstract

Beyond Petroleum (BP), Chevron, ExxonMobil (EM), and Shell, the remaining oil and gas companies of the once largest oil and gas monopoly in the world, Standard Oil, are at risk of losing their positions as major players in the energy industry as the world transitions to a carbon-neutral future by 2050. Investments in low-carbon solutions and alternative energy will be crucial as efforts to diversify energy portfolios accelerate to accommodate a changing economy, especially in renewable energies such as solar, wind, bioenergy, and more. Chevron and EM are the most resistant to this reality as its investments in alternative energy are limited to nonexistent. In contrast, BP and Shell, which have developed major joint-venture projects with other energy companies, have a foundation in establishing themselves as potential energy companies, as opposed to just oil and gas companies.

Keywords: low-carbon, carbon-neutral, renewable energy, fossil fuels, oil and gas companies, greenhouse gas emissions

Introduction

In December 2015, at the United Nations Framework Convention on Climate Change’s (UNFCCC) 21st Convention of Parties (COP 21), 196 parties signed a commitment to limit global temperature increases to 1.5°C compared to pre-industrial levels, known as the Paris Agreement. Now in 2021, humanity is under “code red,” based on the findings of the sixth assessment report released on

August 8th, 2021 by the Intergovernmental Panel on Climate Change (IPCC), a group of scientists working under the United Nations (UN) to provide policymakers with the latest updates on climate change. This report concluded that global temperatures were the highest in 2,000 years in 2020 and are expected to surpass the 1.5°C carbon budget set forth by the Paris Agreement without significant reductions to greenhouse gas emissions (GHGEs) (Masson-Delmotte et al. 2021). Indubitably, this has been the product of anthropogenic activity, with carbon dioxide (CO₂) from the burning of fossil fuels and industrial processes making up two-thirds of GHGEs (United Nations Environment Programme [UNEP] 2021).

Only 100 companies make up 71% of global GHGEs by weight (Griffin 2017). Integrated oil and gas companies (IOGCs) are the principal instigators of increasing GHGEs, starting as early as the Industrial Revolution in the 1800s. Notably, industrialist John D. Rockefeller created Standard Oil, and by capitalizing on the unregulated field through vertical and horizontal integration, he monopolized the oil industry. Standard Oil's success can be attributed to Rockefeller's strict budgetary practices, in which he constantly tried to reduce costs wherever possible in the hopes of long-term benefits, allowing the company to fund itself rather than relying on investors for long-term projects (Pratt 2012). Although Standard Oil dissolved in 1911 after the Supreme Court of the United States ruled that Standard Oil violated the Anti-Sherman Trust Act, its success passed on to its byproducts, known as the seven sisters (Pratt 2012). Today, four out of the seven sisters stand and remain as part of the group of the world's largest IOGCs known as Big Oil: Beyond Petroleum (formerly known as British Petroleum [BP]), Chevron, ExxonMobil (EM), and Royal Dutch/Shell. These four sisters (4S) also stand amongst the top 25 companies responsible for 51% of global GHGEs by weight (Griffin 2017).

U.N. Secretary-General Antonio Guterres has stated that the sixth IPCC report "must sound a death knell for coal and fossil fuels" (Guterres 2021). With temperatures already 1.2°C above pre-industrial levels, climate/carbon neutrality is needed to stay within the 1.5°C carbon budget and to achieve the goals of the Paris Agreement (IPCC 2021). According to the International Energy Agency's (IEA) Stated Policies Scenario (STEPS) in which projections are based on current and announced global policies, renewable energies are expected to dominate the energy sector in the future, with natural gas and oil following behind large margins but projected to still play significant roles in the industry (IEA 2020c). Several factors contribute to this, including the ability to mass-produce components to construct renewable energy projects, and diversifying energy portfolios that support the global energy supply with resources that will not be subjected to future price increases from a lack of supply compared to fossil fuels (Freese et al. 2008).

As clean energy rises in popularity and the climate crisis worsens, the roles of major IOGCs in the energy industry are in question, as has already happened to major coal companies. If the 4S remain loyal to the founding principles of their parent company and continue as IOGCs without expanding into low-carbon/carbon-neutral alternatives or evolving into integrated energy companies (IECs) with diversified energy portfolios, they could be vulnerable to what is known as the Kodak moment, described as the decline of a business caused by failure to adapt to a changing industry and named after Kodak's missed

opportunity to invest in digital photography (Hamilton 2015). In that case, what role, if any, will the 4S play in a carbon-neutral future?

Important Abbreviations

<u>Abbreviation</u>	<u>Meaning</u>
4S	Four sisters
BP	Beyond Petroleum
CCS	Carbon capture and storage
CCUS	Carbon capture, utilization, and storage
CO₂(e)	Carbon dioxide (emissions)
GHGEs	Greenhouse gas emissions
EM	ExxonMobil
GHG Protocol	Greenhouse Gas Protocol
IEC	Integrated energy company
IPIECA	International Petroleum Industry Environmental Conservation Association
IOGC	Integrated oil and gas company

Methodology

According to the Greenhouse Gas Protocol (GHG Protocol), scope 3 emissions are indirect GHGEs that result from the consumption of products sold by a company (GHG Protocol n.d.). They also account for the majority of GHGEs produced by a company by weight (GHG Protocol n.d.). They are organized into 15 sub-categories, but category 11, defined by the GHG Protocol as the use of sold products, is the parameter that each of the 4S has used to report their respective scope 3 emissions (Table 1) due to company concerns including the lack of third-party data and duplicative accounting. Additionally, category 11 is deemed to be the most relevant category to GHGEs reporting for IOGCs by the International Petroleum Industry Environmental Conservation Association (IPIECA 2020). Although the GHG Protocol does not encourage scope 3 emission levels to be compared across companies due to inconsistent reporting behaviors, the data for each of the 4S is nevertheless important to consider to evaluate how sustainable each company is (GHG Protocol).

Table 1. *Estimated scope 3 emissions for each of the 4S and emission reduction commitments/plans (BP 2021e, Chevron 2021d, EM 2021c, Shell 2021b).*

	<u>BP</u>	<u>Chevron</u>	<u>EM**</u>	<u>Shell</u>
Reporting Guidelines Followed	BP Strategy & Sustainability (S&S) carbon management team	IPIECA, Category 11	IPIECA, Category 11	IPIECA, Category 11
2020 Scope 3 Emissions (million tonnes [Mt] CO₂e) <i>The values for each method are not to be aggregated, with the exception of Shell.</i>				
Upstream Production Method	328*	412	540	452
Sales Method	NA***	583	650	
Refining Throughput Method	NA***	372	600	602
Supports the Paris Agreement?	Yes	Yes	Yes	Yes
Net-zero pledge?	Yes, for company operations and the production/sale of products	No	No	Yes, for company operations and the production/sale of products

Note. The upstream production reporting method accounts for scope 3, category 11 emissions that come from the production of sold goods. The sales method accounts for scope 3, category 11 emissions that come from the sale of goods. The refinery reporting method accounts for scope 3, category 11 emissions that come from the processing of energy resources.

* BP's 2020 Sustainability Report states that "its emissions are broadly equivalent to the GHG Protocol, Scope 3, category 11, with the specific scope of upstream production volumes" (BP 2021a), which is the equivalent of IPIECA, scope 3, category 11's upstream production reporting method.

** EM noted its scope 3 emissions to be from the use of specifically sold oil and gas, while it can be implied that other IOGCs account for other energy sources in scope 3 emissions.

*** NA = Not available.

Each of the 4S supports the Paris Agreement’s goal of reaching carbon neutrality in 2050 (Table 1), but the roles that each of them will play to achieve this aim differ. Specifically, Chevron and EM both have not pledged to reach net-zero emissions as a company, which will impact how their support for the Paris Agreement will be proven through their sustainable development plans (Table 1). On the other hand, BP and Shell have both committed to net-zero scope 1, 2, and 3 emissions pledges in terms of reducing company operations and consumer usage of sold products (Table 1).

Given the significance of scope 3 emissions in reaching global net-zero, this research paper will compare the 4S’s plans to reduce their scope 3 emissions transition in response to the transition to a low-carbon/carbon-neutral future through investments in the production and/or distribution of alternative energy to consumers. This includes solar power; hydropower; offshore (water-based generation) and onshore (land-based generation) wind power; geothermal power; bioenergy, including renewable diesel (diesel created entirely from deoxygenized oils/fats), biodiesel (diesel created from a mixture of petroleum diesel and 5% - 20% fuel from oils/fats), sustainable aviation fuels, and biomethane/biogas/renewable natural gas (natural gas produced by the decomposition of waste, typically organic); green hydrogen (created from water undergoing electrolysis and powered by renewable energy sources) and blue hydrogen (created from methane in natural gas) fuel; nuclear power; and electric charging. Indirect GHGE reduction practices irrelevant to sold products and used to help with these processes, such as natural carbon sinks and carbon capture technologies (specifically carbon capture and storage [CCS] or carbon capture, utilization, and storage [CCUS]), as well as hydrogen, will not be considered since they are not sources of energy.

Results

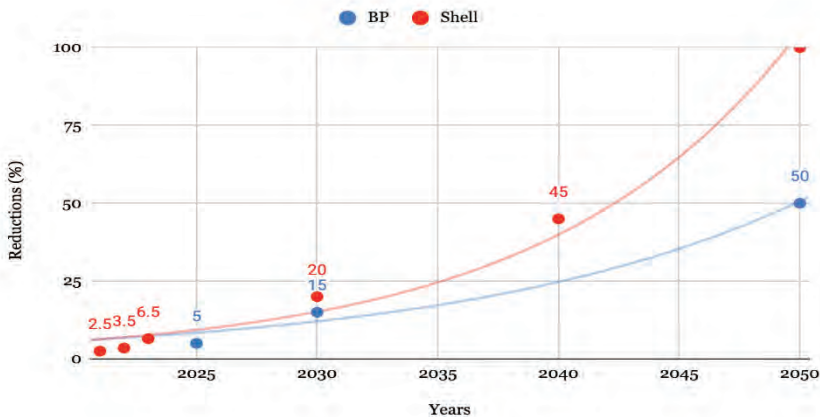


Figure 1. The carbon intensity reduction target benchmarks for sold products for each of the 4S, excluding EM (BP 2021c, Shell 2021).

Reducing the carbon intensity, which is the amount of carbon emissions resulting from the production and/or consumption of energy, will play an important role in achieving a low-carbon/carbon-neutral future. Given that EM has only agreed to report scope 3 emissions starting in 2021 and has not set any targets for carbon intensity reductions, and Chevron has only reported carbon intensity targets for upstream production, both are omitted from Figure 1. Carbon intensity targets for sold products, including emissions from the production, processing, transportation, and end-use of sold products, are reported by BP and Shell. BP aims to achieve 50% reductions by 2050 or sooner, planning to expedite this process during the latter half of years approaching 2050. On the other hand, Shell has the most ambitious plan with its goal to reach 100% reductions by 2050 or sooner.

Table 2. Overview of the types of alternative energy that each of the 4S is investing in for sale (BP 2021c, Chevron 2021c, Shell, 2021c)

Energy Sources		BP	Chevron	EM	Shell
Hydrogen Fuel		Yes	Yes	Yes	Yes
Bioenergy		Yes	Yes	Yes	Yes
Non-bio Renewable Energy	Solar	Yes	Yes	No	Yes
	Wind	Yes	Yes	No	Yes
	Hydro	No	No	No	No
	Geothermal	No	Yes	No	No
Nuclear Power		No	Yes	No	No
Electric Charging		Yes	Yes	No	Yes

Each of the 4S displays varying carbon intensity reduction plans with different means of achieving them, as shown in Table 2. The following results reveal the sustainable development plans of the 4S, including but not limited to joint ventures (a business created/shared by two or more companies), research, investments, virtual power purchase agreements (an agreement that a company sells but is not responsible for the production of energy provided by its supplier), and/or future goals/plans of the 4S to pursue alternative energy.

BP

General Overview of Plans

BP is the only 4S that has explicitly stated a transition from being an oil company to an energy company (BP 2021b). It has developed 20 aims, separated into the three focus areas of its sustainability framework: 1) getting BP and the world to net-zero, 2) caring for the planet, and 3) improving people's lives (BP n.d.-b). The first five are dedicated to BP's company-specific sustainability goals, including: achieving net-zero emissions across operations (scope 1 and 2 emissions) by 2050 or sooner, net-zero across oil and gas upstream productions by 2050 or sooner, halving the carbon intensity of its sold products by 2050 or sooner, reducing

methane intensity in operations to 0.2%, and investing more money into low-carbon/carbon-neutral technologies (BP n.d.-b). In support of BP's fifth aim, its low-carbon investments totaling \$750 million are focused on low-carbon electricity from low-carbon sources such as wind and solar, bioenergy, and hydrogen, in addition to technologies like CCUS (BP 2021b). In terms of oil and gas consumption, BP plans to cut its output of both products by 40% and spend \$5 billion a year on low-carbon projects by 2030 (BP 2021b).

BP Ventures is a venture capital firm founded in 2006 and designed to help BP expand its energy portfolio, specifically through clean energy sources. Since its founding, the firm's purpose has been narrowed down to support BP's net-zero ambition, with all of its ventures aligning with at least one of its five main foci: 1) bio and low-carbon products, 2) carbon management, 3) power and storage, 4) advanced mobility, and 5) digital transformation (BP n.d.-c). Its energy portfolio expansion investments are based on productions and operations, customers and products, gas and low-carbon energy, trading and shipping, and innovation and engineering, in addition to another focus area for new markets (BP n.d.-c). It has invested over \$800 million across more than 60 companies with goals ranging from capturing CO₂ to providing more electrical vehicles with electricity (BP n.d.-c).

Solar

BP's investment plans in solar energy are concentrated in its joint venture, Lightsource BP, a solar development company. Since first joining Lightsource BP in 2016, BP has helped this venture grow its development pipeline from 1.6 GW to 17 GW in regards to its renewable energy generation capacity (BP 2021b). Today, Lightsource BP has a global solar development capacity of 3.5 GW, has powered 1,060,606 homes, and has taken 442,188 cars off the road (Lightsource BP n.d.). However, there are uncertainties about the accuracy of the data (Table 3). The baseline and endline years for these calculations are unspecified; despite Lightsource providing a solar power calculator, its calculator data collection sheet name is only specified with November 2020. Plus, Lightsource BP has data for only its United Kingdom, United States, and Australian sites despite having projects around the world (Lightsource BP 2020). Calculations for the estimated number of homes powered were produced by dividing the estimated annual generation of each site by the average annual energy consumption of a home respective to each site's country (Lightsource BP 2020). The amount of carbon saved was calculated by multiplying estimates of offset emissions embedded at the generation, transmission, and distribution levels by the estimated annual generation of each site (Lightsource BP 2020). To put the amount of carbon saved into perspective, the number of cars that were prevented from being on the road was calculated by dividing the carbon saved in each site by the average amount of CO₂ emissions released by an average car respective to each site's country (Lightsource BP 2020). However, when self-aggregating the data provided, the sums are inconsistent with the website's data (Table 3).

Table 3. Results obtained from Lightsource BP's net solar power calculator (Lightsource BP 2020)

	Estimate Annual Generation (MWh)	Homes Powered	Carbon Saved (tonnes)	Cars off the Road
U.K.	3,000,000	909,091**	849,000	379,018**
U.S.	2,600,000	237,010**	1,838,200	399,609**
Australia	2,600,000	433,333**	2,080,000	693,333**
Website Reported Total	N/A	1,060,606	N/A	442,188
Total*	8,200,000	1,579,434**	4,767,200	1,471,960**

*Self-calculated.

**Rounded to the nearest whole number.

Although Lightsource BP has not released any official reports on its work, its website provides information on its past and current projects. In the United States alone, Lightsource BP has projects in Ohio, Louisiana, Colorado, Texas, Pennsylvania, Indiana, Arkansas, Alabama, Kansas, California, and New Mexico (Lightsource BP n.d.). The renewable energy capacities range from 9 to 375 MW in direct current (dc). Its joint venture partners include, among others, McDonald's, eBay, Verizon, and more (Lightsource BP n.d.). Some of its largest projects in renewable energy capacity include the Birch solar project in partnership with Amazon near Lima, Ohio, which is projected to supply 375 MWdc and offset 423,7000 Mt of CO₂ annually. In addition to this, its Bighorn solar project in Pueblo, Colorado, a 300 MWdc solar farm that was created with Xcel Energy and electric vehicleRAZ North America (an electrical services company and a vertically integrated steel and mining business, respectively), is projected to offset 433,770 Mt of CO₂ annually (Lightsource BP n.d.). The Bighorn project will help Colorado reach its goal of reaching 100% clean energy by 2040 and Xcel's goal to deliver 100% carbon-free electricity by 2050 (Lightsource BP n.d.). Outside the U.S. and highlighted in BP's 2020 sustainability report, one of Lightsource BP's sites in Zaragoza, Vendimia, is set to be a cluster of five solar plants and the company's first major project in Spain (BP 2021b). Once completed, it is expected to give a renewable generating capacity of 246 MW, which would provide clean energy to about 50,000 households (BP 2021b).

Wind

BP's wind energy endeavors have persisted for over ten years through BP Wind Energy and have developed a 9.1 GW generating capacity (BP n.d.-d). It currently has nine onshore wind sites in the U.S., including Colorado, Hawaii, Kansas, Idaho, Indiana, Pennsylvania, South Dakota, and Texas, adding up to a 926 MW generating capacity (BP n.d.-d). In 2018, its wind assets saved about two Mt of CO₂ emissions, although more recent data have not been provided (BP n.d.-d). In January 2021, BP secured a partnership with Equinor, an energy company, to invest \$1.1 billion in additional offshore projects in the U.S. (BP 2021d). Through

this endeavor, 4.4 GW are expected to be added to BP'S generating capacity across four projects by purchasing a 50% interest in Equinor's Empire and Beacon Wind leases on the east coast, marking BP's first investment into offshore wind (BP 2021d).

Bioenergy

BP has not revealed plans to rebrand or redesign its trademarked BP Gasoline and BP Diesel to be more sustainable and is instead focused on efficient consumption (BP n.d-a). However, BP has invested in joint ventures in bioenergy. For example, in 2016, Fulcrum Bioenergy, a company that produces low-carbon fuel from waste, partnered with BP to provide Air BP, the aviation division of BP and one of the largest suppliers of sustainable aviation fuel, with 50 million gallons of low-carbon fuel annually (BP n.d.-a). BP aims to have a 20% share of global sustainable aviation fuel sales by 2030 (BP n.d.). Its main work in bioenergy is concentrated in BP Bunge Bioenergia, which was formed in 2019 as a joint venture between Bunge and BP, focusing on the production of biofuel through the use of sugarcane-derived ethanol (BP Bunge n.d.).

Hydrogen

As part of the Hydrogen Council, a global network of business leaders working to build the hydrogen economy in the energy industry, BP has been working to invest in both green and blue hydrogen (BP 2019). In March 2021, BP announced what would be the U.K.'s largest hydrogen project, in which BP aimed to meet 20% of the U.K.'s hydrogen target by producing up to 1 GW by 2030 through blue hydrogen (BP 2021f). Additionally, with the Australian Renewable Energy Agency, BP plans to use solar energy to power the production of green hydrogen, which could then be used to create green ammonia when combined with nitrogen (BP 2021b). In collaboration with Ørsted, an energy company, BP's first large-scale green hydrogen project will be a 50 MW electrolyzer powered by Ørsted's offshore wind farm (BP 2020).

Electric Charging

BP's CEO of downstream operations, including transportation, distribution, and end-use of sold products, Tufan Erginbilgic, stated that fast and convenient electric charging will be necessary for the widespread adoption of electric vehicles, aligning with BP's integration of electric charging in its strategy for energy transition (BP 2018). Notably, in 2018, BP announced its acquisition of Chargemaster, the U.K.'s largest electric vehicle charging company (BP 2018). The source(s) of the electricity from electric charging is unspecified, however.

Unmentioned Alternative Energies

Geothermal energy, nuclear power, and hydropower were not mentioned in BP's sustainable development plans.

Chevron

General Overview of Plans

In Chevron's 2020 Sustainability Report, the company announced its goal of reaching 24 kg of CO₂ emissions (CO₂e) per barrel of oil equivalent for its oil and natural gas to place itself as one of the most carbon-efficient resources in the world (Chevron 2021c). Part of this goal includes focusing on three aims: 1) achieving lower carbon intensity through cost-efficient methods by investing \$2 billion in carbon-reduction projects, 2) investing \$750 million in renewable energy and carbon offsetting technologies specifically "in support of [its] business;" and 3) investing in low-carbon technologies for consumers (Chevron 2021c). These aims have been notable through its Future Energy Fund, a project series under Chevron Technology Ventures, a subsidiary of Chevron dedicated to exploring and providing clean energy (Chevron n.d.-a). Future Energy Fund I was launched in 2018 and focused on CCUS and energy storage. Taking a slightly different approach with Future Energy Fund II, launched in 2021, Chevron has invested \$300 million to focus more on decarbonizing energy (Chevron 2021b). In terms of scope 3 emissions, Chevron stated that it would address them by supporting a price on carbon, providing transparent data, and enabling customers to lower its emissions (Chevron 2021c).

Solar/Wind Power

Chevron's primary solar and wind power development plans are with Algonquin Power & Utilities Corporation, in which it plans to produce 500 MW of energy through renewables by 2025 (Chevron 2021b). Their initial projects will be based in the U.S. Permian Basin (Texas and New Mexico), Argentina, Kazakhstan, and Western Australia, but only to power Chevron's upstream operations (Chevron 2021b). In 2021, Chevron announced an investment in Ocergy's floating offshore wind platforms but the consumers of this energy are unspecified (Chevron 2021). As such, Chevron currently has not released any plans to distribute and sell solar/wind power to consumers.

Bioenergy

Chevron's investments in bioenergy are focused on various fields including bio and renewable diesel, sustainable aviation fuels, and renewable natural gas. Chevron highlighted that it began to sell diesel fuel containing 6-20% renewable fuel in 2017 (Chevron n.d.). It started selling its own B20 diesel, a mixture of biofuel usually containing 20% biodiesel and 80% petroleum diesel. Its mixture, however, contains up to 20% of biodiesel and renewable diesel each, and "as little as" 60% of petroleum diesel (Chevron n.d.-b). Its current bioenergy projects are focused on producing biofeedstock and renewable natural gas. Its El Segundo Refinery in California is set to produce gasoline, jet fuel, and diesel fuel with biofeedstock to be sold to consumers in Southern California, expectedly by mid-2021 (Chevron n.d.-b). Its joint ventures with both Brightmark LLC and CalBioEnergy produce renewable natural gas for Californian cars and will fund the development and operations of dairy biomethane projects in the

U.S.(Chevron n.d.-b). Chevron has agreed to provide renewable natural gas to California truck operators in July 2020 and sustainable aviation fuels to fuel suppliers, airlines, and airports in 2019, in partnership with Clean Energy Fuels Corporation's Adopt-a-Port initiative and as a part of San Francisco International Airport's agreement to use sustainable aviation fuels (Chevron n.d.-b).

Other Investments

Chevron has highlighted the following investments on its low-carbon technologies page, some of which are part of Future Energy Fund's 15 investments.

- Chevron's progress in developing hydrogen fuel is currently limited. As part of the California Fuel Cell Partnership, the Hydrogen Council, Oil and Gas Climate Initiative (OGCI), work is being conducted to explore the possibilities of hydrogen fuel and eventually provide it to consumers (Chevron n.d.-a).
- Chevron has invested in Zap Energy, a startup company focused on producing energy through nuclear fusion. The invested amount is unspecified, as well as if the energy will be for company operations or if the investment is a virtual power purchase agreement for energy to be sold to consumers (Chevron n.d.-a).
- Chevron has invested in Eavor Technologies, a global geothermal technology company. Similar to its investment in Zap Energy, the amount invested and the reason for investment were not unspecified (Chevron n.d.-a).

Unmentioned Alternative Energies

Hydropower and electric charging were both unmentioned in Chevron's sustainable development plans.

EM

EM created EM Low-Carbon Solutions to explore and implement carbon-reducing technologies. Its main foci reside in carbon capture and storage (CCS) with hydrogen as opposed to CCUS, advanced biofuels, lightweight packaging, and efficient lubricants and fuels (EM 2021c). Since 2000, EM has invested over \$10 billion in CO₂ reduction methods (EM n.d.-b).

Solar/Wind Energy

On its frequently asked questions page, EM acknowledges that renewable energy including solar and wind is important in transitioning to the use of low-carbon energy (EM n.d.-b). It mentions that it purchased 500 MW of energy through wind and solar power in Texas, but fails to mention that the energy will be used to power crude oil production (EM n.d.-b). EM cites the IEA's Tracking Clean Energy Progress report, stating that "the expansion of wind and solar will not be sufficient to achieve the goals of the Paris Agreement" (EM n.d.-b). However, the

IEA report states that these investments are not on track, not necessarily that they are not enough to reach a low-carbon/carbon-neutral future (IEA 2020b). Instead of investing in renewables, EM is focusing on new “breakout technologies” through its research and development team in areas with limited methods to large-scale CO₂ reductions, including commercial transportation, power generation, and industrial processes (EM n.d.-b)

Bioenergy

In the past decade, EM has invested \$250 million in biofuels research. Some of its projects in bioenergy include the following:

- In a partnership with Synthetic Genomics Inc, EM has been working on the production of biofuel by using algae and researching the use of algae and plant waste (EM n.d.-a). Together, their goal is to produce 10,000 barrels of algae biofuels by 2025 (EM n.d.-a).
- Similarly, EM and Renewable Energy Group have also partnered with Clariant, a specialty chemicals company, on research to produce biofuel from agricultural waste and residue (EM n.d.-a).
- EM agreed to purchase 2.5 million barrels of renewable diesel annually for five years from Global Clean Energy Holdings starting 2020 (EM n.d.-a)
- EM has also completed a trial of its first bio-based marine fuel, which it claims will provide a 40% reduction in CO₂e compared to traditional marine fuel (EM n.d. -a).

Unmentioned Alternative Energies

Geothermal energy, nuclear power, hydropower, hydrogen (specifically as fuel to be sold to consumers; hydrogen for CCS technology was mentioned but will not be considered), and electric charging were all unmentioned in EM’s sustainable development plans.

Shell

General Overview of Plans

Shell aims to be a net-zero emissions company by 2050 by reducing all of its scope 1, 2, and 3 emissions, specifically noting that this also means using CCS technology and balancing emissions with offsets (Shell 2021b). It acknowledges that 90% of its emissions come from the products it sells, including the products it does not produce; thus the company is working to decarbonize electricity through increasing operational efficiency in company assets, shifting to natural gas, growing its low-carbon power business, and providing low-carbon solutions including biofuels and hydrogen, CCS, and nature-based offsets/natural sinks (Shell 2021b). As such, in consideration of its scope 3 emissions, Shell has committed to reducing the carbon intensity of the products it sells by 100% by 2050 (Table 2).

All of Shell’s electricity is 100% renewable and is certified by Renewable Energy Guarantees of Origin issued by Ofgem, the energy regulator of Britain,

meaning that with every unit of energy used, a unit of renewable energy is produced in the U.K. by renewable sources including wind, solar, and biomass (Shell n.d.-b). However, this does not guarantee that the energy directly used is renewable and rather “is a blend of electricity generated from all the different sources available—wind, solar, gas, nuclear, coal, etc.” (Shell n.d.-b). Shell’s Renewables and Energy Solutions, previously known as New Energies, was created in 2016 to explore low-carbon solutions (Shell 2021b). With the help of Shell Ventures, Shell’s venture capital firm founded in 1996, Shell is able to invest in new technologies that complement its business.

Solar

Shell listed various investments in solar energy in its 2020 Energy Transition Strategy report, but does not specify if the energy produced will be used for company operations or distributed to consumers (Shell 2021b). However, in its 2020 sustainability report, it explains its solar energy investment plans in more detail. Specifically, by the end of 2020, its solar generating capacity was up to 674 MW, with an additional 1,053 to come (Shell 2021b). In the report, it highlights a few solar power projects, including its company, Silicon Ranch, U.S., that had a generating capacity of 1,043 MW by the end of 2020, and the Qabas project in Oman that provides 25 MW to a smelting company annually (Shell 2021b). Additionally, Shell states that it has made minority investments in companies to provide clean and reliable energy to consumers, such as PowerGen, a company that builds and manages solar grids in Africa, and d.light, a company that provides solar power to companies and homes across 70 countries (Shell 2021b).

Wind

Shell has over 6 GW of energy from its wind projects. Its involvement is primarily in offshore wind, which started in 2000 (Shell n.d.-c). Some of its offshore projects include its joint ventures Atlantic Shores and Noordzeewind, which will produce 2.5 GW in New Jersey and produce over 100 MW in the Ditch North Sea (Shell n.d.-c). It is also part of the Blauwwind and Mayflower consortiums, whose projects produce 731.5 MW off the Dutch coast and will produce 804 MW off the Massachusetts coast (Shell n.d.-c). In terms of new technologies, Shell has invested in floating and kite wind turbines with companies including Makani and Simply Blue Energy to resource wind energy in additional areas (Shell n.d.-c). In U.S. markets, Shell’s involvement in onshore wind first started in 2001, and it currently operates four onshore wind farms, one each in Wyoming and Texas along with two in California (Shell n.d.-c).

Hydrogen

Shell has invested in hydrogen but “will begin by producing and supplying hydrogen for [its] own manufacturing sites, especially refineries” (Shell 2021b). The type of hydrogen fuel (green or blue) that Shell has invested in is unspecified.

Bioenergy

Shell's work in bioenergy is largely concentrated in sustainable aviation fuels. It is a founding partner of the Clean Skies for Tomorrow Coalition, a group launched by the World Economic Forum and Energy Transitions Commission consisting of airlines, airports, fuel providers, and engine manufacturers to reduce GHGs in the aviation industry by promoting the widespread use of sustainable aviation fuels (Shell 2021b). In fact, one of Shell's main partnerships is its century-long collaboration with Rolls Royce to test new types of technology, fuels, and infrastructure in commercial aviation (Shell 2021b). Today, Shell is testing the capability of Rolls Royce engines to run 100% on sustainable aviation fuels (Shell 2021b).

Aside from sustainable aviation fuels, another important project of Shell's is Raízen, its joint venture with Cosan, a Brazilian energy company that focuses on bioethanol, sugar, and energy, to produce low-carbon biofuel from sugarcane in Brazil (Raízen n.d.). Raízen is the first company to produce secondhand ethanol and is the largest producer of energy from biomass in Brazil (Raízen n.d.). Raízen has also invested in its first plant to produce biogas in 2020 by reusing its industrial residue during operations (Raízen n.d.).

Electric Charging

Shell has invested in various new electric charging technologies. This includes providing more electric charging home stations through its NewMotion company, which is also "one of Europe's largest electric-vehicle charging providers" (Shell n.d.-a). Additionally, Shell plans to install 50,000 street electric charging stations in the U.K. by 2050 as a complement to its home charging station initiatives (Shell n.d.-a). Both of these projects are part of Shell's plan to accommodate electric vehicle drivers in hopes of increasing the popularity of using electric vehicles (Shell n.d.-a).

Unmentioned Alternative Energies

Nuclear power, geothermal power, and hydropower were all unmentioned in Chevron's sustainable development plans.

Discussion

General Observations

Data Collection

Given the drastically varying methods of reporting information and statistics, quantitative comparisons were difficult to make amongst the sustainable development plans of the 4S. For example, BP was the only company that specified what type of unit it measured its solar generating capacity in by providing conversions in consideration of dc and alternating currents (AC), although it highlighted the measurements in the overviews of its solar projects in

dc voltage since DC voltages present higher values. Similarly, for projections and recorded results for solar and wind production, the generating capacity was measured in the number of houses powered. This was common across the data provided by the 4S, but these metrics were not considered because of missing information that would specify the methods used to calculate them, and the data were consequently unable to be compared on a consistent basis. Additionally, information about joint ventures, investments, and partnerships generally follow one of two patterns: (1) the information discusses what the partnering companies have done in the past and/or (2) the expectations of what the project will achieve. As such, progress reports with results for most joint ventures were either difficult to find or nonexistent.

The Future of Oil and Gas

Unsurprisingly, information about the future of oil and gas production was widely unavailable. This can be due to numerous reasons including investor, board, or company concerns of appearing hypocritical in terms of being sustainable by continuing oil and gas production while investing in low-carbon solutions/alternative energy sources. In fact, information regarding oil and gas production was based largely on reductions for upstream emissions (i.e. scope 1 and 2 emissions), including reducing methane emissions and using carbon capture technologies to do so. None of the 4S have committed to phasing out oil and gas entirely and continue to see both energy sources as major players in the energy industry in 2050.

The Lack of Hydropower Investments

According to the IEA, hydropower is the largest renewable energy source in the world in terms of energy generation (IEA 2020). Regardless, none of the 4S have invested in hydropower or even mentioned it in their sustainable development plans, nevermind reasons as to why they decided to invest in other energy sources besides hydropower. This may be because the production process of hydropower is drastically different from those of solar and wind given its reliance on topography and heavy usage of concrete (which is carbon-intensive) for construction. Plus, hydropower is consequently viewed separately from solar and wind, which can and are often grouped together in discussion.

Commitment to a Carbon-Neutral Future in 2050

All of the 4S have explicitly stated their support for the Paris Agreement. Regardless of whether their work aligns with global net-neutrality, company net-neutrality, or both, all of the companies have made efforts to reduce GHG emissions. However, even with the hundreds of millions of dollars of investments made, these expenses account for a small percentage of the billions of dollars in profit that the 4S each make annually. In fact, the IEA found that the industry only spent 1% of its capital expenditures (funds used to buy or improve a company's fixed assets) on clean energy in 2020 (IEA 2021). With the 4S in 2020 alone, BP spent \$12 billion in capital expenditures; Chevron spent \$9 billion in capital expenditures in 2020, yet its Future Energy Fund project has only been invested

with \$300 million; EM spent \$21 billion in capital and exploration expenditures but only focused \$250 million on bioenergy research, its main low-carbon focus; and Shell spent \$16.5 billion in capital expenditures, with \$907 million of it going to research and development (BP 2021b; Chevron 2021e; EM 2021b; Shell 2021c). As such, the authenticity of the 4S's support for not only the Paris Agreement but also the possibility of a carbon-neutral future and investing in alternative energy can be challenged.

American versus European IOGCs

BP is based in the U.K. and Shell is based in the Netherlands, whereas both Chevron and EM are based in the U.S. The correlation between company location and company sustainable development plans is apparent, considering that BP and Shell tend to invest in energy sources while Chevron and EM have invested in new technologies to sustain their production of oil and gas. Specifically, BP and Shell are the only companies that have set carbon-intensity reduction targets for sold products and have already heavily invested in solar/wind power compared to Chevron and EM, which have done neither. The U.S. notably backed out of the Paris Agreement in 2020 and only recently rejoined it in 2021, demonstrating a relative lack of priority for environmental issues in the government, which is reflected in Chevron and EM's sustainable development plans.

BP

For the most part, BP has remodeled its company through its 20 aims and three focus areas for sustainability. However, despite having goals of becoming a net-zero emissions company by 2050, its carbon intensity goals for the products it sells (Figure 2) may prevent them from doing so. BP's second aim implies that it plans to continue selling oil and gas despite its increasing investments in low-carbon solutions. In fact, the majority of the aims shift the responsibility of the current climate crisis to consumers, with only one-fourth of BP's aims being relevant to the company itself. BP's plentiful investments and joint ventures in low-carbon solutions that already have a foundation demonstrate a stronger representation of BP's commitment to sustainability. BP's hydrogen investments are also the most established, with EM, Chevron, and Shell all currently planning to use hydrogen for CCS/CCUS technology for now, whereas BP has been producing energy through hydrogen.

Chevron

Chevron appears resistant to investing in clean energy, especially as the company has not released plans to sell solar and wind power to consumers. Solar and wind energy are the fastest-growing renewables (IEA 2020), but Chevron has yet to invest in them commercially. The bulk of Chevron's sustainable development plans are through its Future Energy Fund project, yet many of its ventures appear to be in development. As briefly mentioned, detailed information about the ventures themselves is unavailable. Unlike the other 4S, Chevron has invested in nuclear fusion power and geothermal energy, yet both technologies are still in development. Especially considering that none of the other 4S have invested in

these two energy sources specifically, it begs the question of whether or not Chevron is purposely investing in low-carbon solutions that are relatively unestablished and will take a long period of time to fully develop, implement, and prove to be efficient sources of energy.

EM

Of all the 4S, EM appears to be the most resistant to a changing energy industry. Its reluctance to renewable energies aligns with its views of achieving the goals of the Paris Agreement as aspirational (Table 1). Like Chevron, EM has not announced any plans to invest in solar/wind power to sell to consumers. Its investment in algae biofuel, in particular, a relatively new and unexplored industry, demonstrates a form of “greenwashing,” similar to Chevron’s efforts to invest in geothermal energy but specifically with nuclear fusion—a difficult yet net-zero emissions energy source to produce. EM’s main foci, however, are on carbon capture, with its work in the field capturing 20% of global GHGs and on biofuel (EM 2021c). In response to a question regarding other low-carbon solutions besides CCS, EM includes plans to reduce methane flaring and operational emissions, emissions that only pertain to scope 1 and 2 (EM n.d.-b). Unless EM’s investors urge the company to pursue alternative energy, EM will most likely remain content as a major IOGC.

Shell

Shell, along with BP, appears to be the most committed to clean energy out of the 4S. It is the only company that acknowledges that most of its emissions are scope 3 and thus require greener products to be sold to reduce emissions. However, part of its decarbonization strategy is to invest more in natural gas (Shell 2021b); although natural gas is a less carbon-intensive energy source than oil, carbon reductions would be more efficiently achieved through its other low-carbon solution investments. Unlike the other companies, Shell has focused on selling wind energy and bioenergy to consumers. Its commitment to 100% renewable energy is also unique; an increase in energy consumption correlates to an increase in renewable energy production in alignment with its renewable energy plan, but since the energy that Shell provides to its consumers is mixed, it also increases the consumption of oil and gas. As such, the emissions produced from this will still remain in the atmosphere despite the production of clean energy along with it. With all things considered, however, Shell’s renewable energy commitment is unmatched in comparison to the other 4S.

Conclusion

Even if each of the 4S became carbon-free companies tomorrow, it would not be enough to reach the goals of the Paris Agreement; hundreds of years of GHGs remain in the atmosphere and consumers still rely on IOGCs for energy demands. Major IOGCs are responsible for the majority of global GHGs, but the global networks of consumers, investors, and political leaders also play roles in the abilities of IOGCs to transition to cleaner energy sources through their willingness

to switch to them. The oil and gas ecosystem has intertwined itself into the global political sphere, affecting the environmental regulations implemented and reinforced; the economic sphere in a largely capitalistic world that has prioritized profit over environmental concerns; and the social sphere, where climate anxiety has begun to spread across populations for fear of catastrophic climate disasters. As such, BP, Chevron, EM, and Shell have the opportunity to set precedents in the oil and gas industry. Given the dire need for cleaner energy, however, the 4S can either serve as leading examples for other IOGCS through their efforts to transition to a carbon-neutral future or as reminders of a once carbon-intensive past.

References

- BP. (n.d.-a) *Fuels and lubricants*. Retrieved from https://www.bp.com/en_us/united-states/home/who-we-are/what-we-do/fuels-and-lubricants.html
- BP. (n.d.-b) *Our aims*. Retrieved from <https://www.bp.com/en/global/corporate/who-we-are/our-ambition/our-aims.html>
- BP. (n.d.-c) *Our portfolio*. Retrieved from https://www.bp.com/en/global/bp-ventures/our-portfolio.html#tab_new-markets
- BP. (n.d.-d) *Wind energy sites*. Retrieved from https://www.bp.com/en_us/united-states/home/where-we-operate/wind-energy-sites.html
- BP. (2018, June 28). *BP to acquire the UK's largest electric vehicle charging company*. Retrieved from <https://www.bp.com/en/global/corporate/news-and-insights/press-releases/bp-to-acquire-uks-largest-electric-vehicle-charging-company.html>
- BP. (2019, June 11). *BP joins the Hydrogen Council*. Retrieved from <https://www.bp.com/en/global/corporate/news-and-insights/reimagining-energy/bp-joins-hydrogen-council.html>
- BP. (2020, November 10). *bp and Ørsted to create renewable hydrogen partnership in Germany*. Retrieved from <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/news-and-insights/press-releases/bp-and-orsted-to-create-renewable-hydrogen-partnership-in-germany.pdf>
- BP. (2021a). *BP annual report and form 20-F 2020*. Retrieved from https://www.annualreports.com/HostedData/AnnualReports/PDF/LS_E_BP_2020.pdf
- BP. (2021b). *Fourth quarter and full year 2020*. Retrieved from <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/investors/bp-fourth-quarter-2020-results.pdf>
- BP. (2021c). *Sustainability report 2020*. Retrieved from <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/sustainability/group-reports/bp-sustainability-report-2020.pdf>
- BP. (2021d, January 29). *bp completes entry into offshore wind, with strategic partner Equinor*. Retrieved from https://www.bp.com/en_us/united-states/home/news/press-releases/bp-completes-entry-into-offshore-wind-with-strategic-partner-equinor.html

- BP. (2021e, March). *ESG Datasheet 2020*. Retrieved from <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/sustainability/group-reports/bp-esg-datasheet-2020.pdf>
- BP. (2021f, March 18). *bp plans UK's largest hydrogen project*. Retrieved from <https://www.bp.com/en/global/corporate/news-and-insights/press-releases/bp-plans-uks-largest-hydrogen-project.html>
- BP Bunge. (n.d.) *About us*. Retrieved from <https://bpbunge.com.br/en/about-us/>
- Chevron. (n.d.-a). *Future Energy Fund overview*. Retrieved from <https://www.chevron.com//media/chevron/technology/documents/future-energy-fund-overview.pdf>
- Chevron. (n.d.-b). *Growing lower carbon businesses*. Retrieved from <https://www.chevron.com/sustainability/environment/growing-lower-carbon-businesses>
- Chevron. (2021a). *Climate change resilience report*. Retrieved from <https://www.chevron.com/-/media/chevron/sustainability/documents/climate-change-resilience-report.pdf>
- Chevron. (2021b, February 25). *Chevron commits \$300 million toward low-carbon technology investments*. Retrieved from <https://www.chevron.com/stories/chevron-commits-300-million-toward-low-carbon-technology-investments>
- Chevron. (2021c). *2020 Corporate sustainability report*. Retrieved from <https://www.chevron.com/-/media/shared-media/documents/chevron-sustainability-report-2020.pdf>
- Chevron. (2021d). *2020 Performance data*. Retrieved from <https://www.chevron.com/-/media/shared-media/documents/2020-sustainability-performance-data.pdf>
- Chevron. (2021e). *2020 Supplement to the annual report*. Retrieved from <https://www.chevron.com/-/media/shared-media/documents/2020-chevron-annual-report-supplement.pdf>
- ExxonMobil. (n.d.-a) *Advanced biofuels*. Retrieved from <https://corporate.exxonmobil.com/Energy-and-innovation/Advanced-biofuels>
- ExxonMobil (n.d.-b) *Frequently asked questions*. Retrieved from <https://corporate.exxonmobil.com/Sustainability/Emissions-and-climate/Frequently-asked-questions>
- ExxonMobil. (2021a). *Sustainability report highlights*. Retrieved from <https://corporate.exxonmobil.com/-/media/Global/Files/sustainability-report/publication/Sustainability-Report.pdf>
- ExxonMobil. (2021b, February 2). *ExxonMobil reports results for fourth quarter 2020 and provides perspective on forward plans*. Retrieved from <https://corporate.exxonmobil.com/-/media/Global/Files/investor-relations/quarterly-earnings/earnings-announcements/2020-earnings-announcements/4q-earnings-release.pdf>
- ExxonMobil. (2021c, April). *Updated 2021 energy & carbon summary*. Retrieved from <https://corporate.exxonmobil.com/-/media/Global/Files/energy-and-carbon-summary/Energy-and-Carbon-Summary.pdf>
- Freese, B., Clemmer, S., & Noguee, A. (2008, October). *Coal Power in a Warming World*. Union of Concerned Scientists. Retrieved from <https://www.ucsusa.org/sites/default/files/2019-09/Coal-power-in-a-warming-world.pdf>

- Greenhouse Gas Protocol. (n.d.). *Corporate value chain (scope 3) accounting and reporting standard*. Retrieved from https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporting-Standard_041613_2.pdf
- Griffin, P. (2017, June). *The Carbon Majors Database: CDP Carbon Majors Report 2017*. CDP. Retrieved from <https://climateaccountability.org/pdf/CarbonMajorsRpt2017%20Jul17.pdf>
- Guterres, António. (2021, August 9). *Secretary-General's statement on the IPCC Working Group 1 report on the physical science basis of the sixth assessment*. United Nations. Retrieved from <https://www.un.org/sg/en/content/secretary-generals-statement-the-ipcc-working-group-1-report-the-physical-science-basis-of-the-sixth-assessment>
- Hamilton, T. (2015). *Die or Adapt: Can Big Oil Transition to a Low-Carbon Economy?* *Corporate Knights*, 14(4), 27-29. Retrieved from <http://www.jstor.org/stable/44149897>
- IEA. (2020a). *Global energy review 2020*. IEA. Paris. Retrieved from <https://www.iea.org/reports/global-energy-review-2020>
- IEA. (2020b). *Tracking power 2020*. IEA. Paris. Retrieved from <https://www.iea.org/reports/tracking-power-2020>
- IEA. (2020c). *World energy outlook 2020*. IEA. Paris. Retrieved from <https://www.iea.org/reports/world-energy-outlook-2020>
- IEA. (2021). *World energy investment 2021*. IEA. Paris. Retrieved from <https://www.iea.org/reports/world-energy-investment-2021>
- IPCC. (2021). *Working Group I contribution to the sixth assessment report of the Intergovernmental Panel on Climate Change*. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM.pdf
- IPIECA. (2016, June). *Estimating petroleum industry value chain (Scope 3) greenhouse gas emissions. Overview of methodologies*. Retrieved from <https://www.iecea.org/resources/good-practice/estimating-petroleum-industry-value-chain-scope-3-greenhouse-gas-emissions-overview-of-methodologies/>
- Lightsource BP. (n.d.). *Energy*. Retrieved from <https://www.lightsourcebp.com/sustainability/energy/>
- Lightsource BP. (2020, November). *Solar power calculations for U.S., U.K., and Australia*. Retrieved from <https://s3-eu-west-1.amazonaws.com/assets.lightsource-re.com/2020/11/Solar-power-calculations-sheet-for-UK-US-and-Australia-112020.xlsx>
- Pratt, J. (2012). *Exxon and the Control of Oil*. *The Journal of American History*, 99(1), 145-154. Retrieved from <http://www.jstor.org/stable/41510310>
- Raízen. (n.d.). *Renewables*. <https://www.raizen.com.br/en/our-business/renewables#ethanol>
- Shell. (2021a). *Outlook for 2021*. Retrieved from <https://reports.shell.com/annual-report/2020/strategic-report/strategy-business-and-market-overview/strategy-and-outlook/outlook-for-2021-and-beyond.php>
- Shell. (2021b). *Shell energy transition strategy 2021*. Retrieved from <https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/sustainability/group-reports/bp-sustainability-report-2020.pdf>

- Shell. (2021c). *Shell sustainability report 2020*. Retrieved from <https://reports.shell.com/sustainability-report/2020/servicepages/downloads/fileshell-sustainability-report-2020.pdf>
- Shell. (n.d.-a). *Electric vehicle charging*. Retrieved from <https://www.shell.com/energy-and-innovation/new-energies/electric-vehicle-charging.html>
- Shell. (n.d.-b). *100% renewable electricity as standard*. Retrieved from <https://www.shellenergy.co.uk/energy/renewables>
- Shell. (n.d.-c). *Wind power*. Retrieved from <https://www.shell.com/energy-and-innovation/new-energies/wind.html>
- Systemiq. (2020, December). *The Paris effect: How the climate agreement is reshaping the global economy: Executive summary*. Retrieved from https://www.systemiq.earth/wp-content/uploads/2020/12/The-Paris-Effect_SYSTEMIQ_Executive-Summary_December-2020.pdf
- United Nations Environment Programme. (2021). *Making peace with nature: A scientific blueprint to tackle the climate, biodiversity and pollution emergencies*. Retrieved from <https://www.unep.org/resources/making-peace-nature>



Putting Constraints on the Dark Matter Density at the Center of our Galaxy from the Most Recent Observations of GRAVITY

Efe Uğurlu

Author Background: *Efe Uğurlu grew up in Turkey and currently attends Robert College in Istanbul, Turkey. His Pioneer research concentration was in the field of physics/astronomy and titled “Relativity, Cosmology, and the Contents of the Universe.”*

Abstract

By investigating the orbital dynamics of stars orbiting the Sgr A* supermassive black hole, we hypothesize the presence of a dark matter wind due to the gravitational pull of the Sgr A*. The wind that is radially flowing into the black hole interacts with the stars that orbit the black hole. During this interaction, the momenta of dark matter particles are transferred to the stars, which causes a radial force on the stars towards the black hole. This force has observable effects on the orbital dynamics of the stars, such as a perihelion shift. In this paper, we derive a theory that describes the perihelion shifts of stars due to the dark matter as a function of distance and put constraints on the dark matter density at the center of our universe using the perihelion shift data of S2 that is gathered via the GRAVITY instrument of the Very Large Telescope Interferometer (VLTI). We also make predictions for other stars and comment on which stars would yield significantly observable results due to the predicted effect. Further studies in this area of our galaxy could make these constraints more accurate.

1. Introduction

1.1 Dark Matter

Dark matter, which constitutes 27% of our universe, is much more abundant than normal matter which constitutes only 5% of our universe [1]. Yet, its existence was not even speculated until the 20th century. One of the earliest studies of dark matter was in 1933; astronomer Fritz Zwicky discovered that the light matter in the Coma cluster was not enough to keep the cluster together [2]. According to his work, the matter in the cluster should have been about 100 times more than the observed matter in order to provide the necessary

gravitational force. He speculated that there must have been more matter in the universe, which he called dark matter, but he could never find what that matter actually was. The mystery of dark matter continued for decades, and it is still continuing today. We know more about what dark matter is not than what it is [3]. There have been many candidates for what dark matter is, such as massive compact halo objects (MACHOs) and weakly interacting massive particles (WIMPs) [4]. While some scientists looked for answers in the quantum world and thought that quantum particles called axions could be dark matter, others thought that the dark matter particles were proton-size black holes called primordial black holes [4]. Many theories have been ruled out, and the remaining ones are not wholly sufficient. Even though there is no actual evidence for what dark matter is, its existence is almost undeniable due to the abundance of observational evidence.

1.1.1 Rotation Curves

Rotation curves of galaxies are one of the most compelling pieces of evidence for the existence of dark matter. These curves are obtained by drawing a velocity versus radius graph for galaxies. If light-matter (light-emitting matter) was the only type of matter that was present in our universe, then the velocity of stars and dust traversing the galaxy would start to decrease gradually after a certain radius [5]. However, this expected decrease either occurs less often than expected or never occurs. One of the studied galaxies, the M33, is a great example of this phenomenon. A decrease in velocity of the dust traversing the M33 galaxy is expected after about 3 kiloparsecs from the galactic center, however, this decrease never occurs. Instead, the observed speed of dust is more than twice the expected speed at 10 kiloparsecs. The most reasonable explanation for this unexpected observation is that there must be more matter in our universe that does not seem to interact with light: in other words, dark matter.

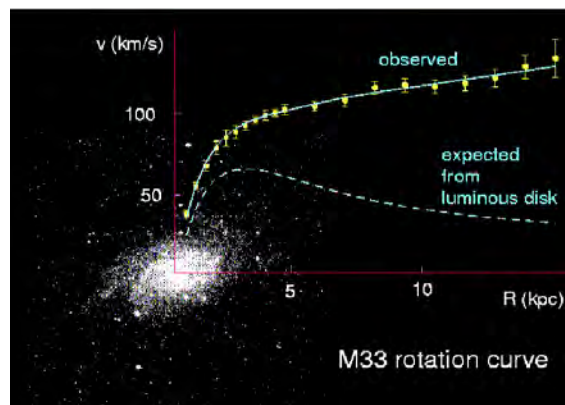


Figure 1. The observed and expected rotation curves of the M33 spiral galaxy [6]

1.1.2 Gravitational Lensing

Gravitational lensing is additional cogent evidence for the existence of dark matter. Gravitational lensing is the phenomenon that occurs when two bright astronomical objects align. As general relativity predicts, light bends around heavy objects, thus the light that comes from the object that is further away from us gets bent by the closer object. This not only enables us to see the bright object that is hidden behind the other bright object but also allows us to see the bright object in the back at multiple different places on the sky [5]. From these images of the object in the sky, we can measure the deflection of light, allowing us to measure the mass of the closer bright object. The measured mass for the closer bright object is, however, much greater than its actual mass. This indicates that there must be additional matter that interacts with light gravitationally, causing it to bend. Mordecai Milgram's Modified Newtonian Dynamics (MOND), another theory that tries to explain these observations, does not predict such a phenomenon, which makes theories for the existence of dark matter much more popular [7].

1.2 Interactions between Normal and Dark Matter

It is almost undeniable that dark matter interacts with normal matter and light, but the exact nature of this interaction is yet to be uncovered. Some scientists speculated the existence of a multiverse and suggested that dark matter is just normal matter in parallel universes, and it interacts with normal matter in our universe only in the form of gravitational waves. In this paper, we speculate that dark matter can be absorbed by large stars. This is likely to occur around black holes because the gravitational pull of the black hole can cause dark matter to flow into it. The flowing dark matter wind would most probably collide with large stars orbiting the black hole and be absorbed by them, transferring its momentum to the stars during the process. This interaction could have observable effects on the orbital dynamics of the star, such as perihelion precession. Recently, careful measurements have been made by the GRAVITY instrument of VLTI for a star named S2 that is orbiting the central black hole of our galaxy. If the hypothesized effect is valid, it can help us understand the nature of dark matter at the center of our galaxy and possibly even put constraints on some of its properties. Testing this effect and obtaining results for the dark matter at the center of our galaxy would also be beneficial and intriguing since most of the measurements for dark matter have been made in our solar system. In this paper, we are going to put constraints on the density of dark matter at the center of our universe by utilizing the perihelion shift data of S2 from the GRAVITY.

2. Dark Matter Force

As stated in the introduction, this paper assumes the presence of dark matter particles flowing into the central black hole of our galaxy. The dark matter particles can be absorbed by S2, and during this interaction, the momenta of the dark matter particles are also absorbed. This change in the momentum of S2 results in a force (F_{DM}) being applied to S2 in the r-direction (in spherical

coordinates) towards the Sgr A*. In this section, this dark matter force as a function of r , the distance between S2 and the Sgr A* black hole, will be derived.

Force, by definition, is the derivative of momentum:

$$F_{DM} = \lim_{\Delta \rightarrow \infty} \frac{\Delta P_{DM}}{\Delta t} \quad (1)$$

Momentum is the total mass times the velocity:

$$P_{DM} = N_{DM} m_{DM} v_{DM} \quad (2)$$

where N indicates the total number of dark matter particles, m is the mass of a single particle, and v is the velocity of each particle relative to S2. We can define a new variable (Γ), which represents the number of particles per unit time:

$$\Gamma = \frac{N}{\Delta t} \quad (3)$$

Now we can represent the force as

$$F_{DM} = n_{DM} v_{DM} \Gamma_{DM} \quad (4)$$

This new variable, Γ , can be found by imagining a set of particles flowing through a surface. Flux, by definition, is the number of particles that go through a surface with area A in time Δt , and it can also be written as number density of particles (n) times their velocity:

$$\Phi = \frac{N}{\Delta t A} = n v \quad (5)$$

Thus, Γ can be written in terms of flux as

$$\Gamma = \Phi A = n v A \quad (6)$$

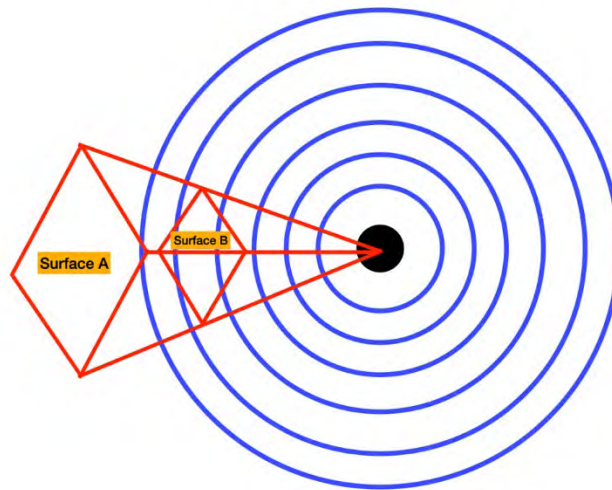


Figure 2. The black circle at the center represents the Sgr A* supermassive black hole. Each blue circle represents a ring of dark matter, and the distance between each circle is infinitesimal. Surface A and surface B are not rectangles but are bent surfaces, parallel to the dark matter rings. The amount of dark matter in each ring is the same, and the rings are moving towards the black hole with a speed that is equal to their escape velocity.

Plugging in equation 6 to equation 4 gives us our final equation for the dark matter force:

$$F_{DM} = v_{DM} m_{DM} v_{DM}^2 A_{S2} \quad (7)$$

The velocity of the dark matter particles is accepted to be equal to their escape velocity since the dark matter particles that are very far away from the black hole are accepted to be stationary:

$$v_{DM} = \sqrt{\frac{2GM}{r}} \quad (8)$$

where G is the gravitational constant, M is the mass of Sgr A*, and r is the distance from the center of the black hole. The area of S2 that is interacting with the dark matter can be naively accepted as πR_{S2}^2 . Also, the number density of dark matter multiplied by the mass of each particle is equal to the density of dark matter. The force equation can be rewritten as

$$F_{DM} = \rho_{DM} v_{DM}^2 A_{S2} = \frac{2\rho_{DM} GM \pi R_{S2}^2}{r} \quad (9)$$

The only unknown variable in this force equation is the density of the dark matter, which is dependent on the distance from the center of the black hole, r :

$$\rho_{DM} = \frac{\eta}{r^a} \quad (10)$$

where η is an unknown constant. The power of r (a) can be determined by noticing that the flux through surface A should be equal to the flux through surface B (Figure 2). This means that the flux is constant:

$$\Phi_{DM} = n_{DM}v_{DM} = \text{const.} \quad (11)$$

Since the mass of each particle is constant, the density of dark matter multiplied by the velocity of dark matter should also be constant:

$$\Phi_{DM}m_{DM} = n_{DM}m_{DM}v_{DM} = \rho_{DM}v_{DM} = \text{const.} \quad (12)$$

This means that the power of r in ρ_{dm} should be $\frac{1}{2}$ since the power of r in v_{DM} is $-\frac{1}{2}$. Now, we can change a in equation 10 to $-\frac{1}{2}$ and plug that expression into equation 9:

$$F_{DM} = \frac{2GM\pi\eta R_{S2}^2}{\sqrt{r}} \quad (13)$$

3. The Effect of Dark Matter Wind on the Orbit of S2

In the previous section, we derived the equation for the force on stars such as S2 that is created by dark matter wind flowing radially to the central black hole. In this section, we will observe the effects of this force on the orbit of the stars and derive an equation for the precession of S2 due to this force. This requires utilizing the Effective Theories to make a correction to the Newtonian orbital dynamics. In order to do this, first, the potential energy due to the dark matter wind must be found.

Force, by definition, is equal to the negative derivative of potential energy:

$$F_{DM} = -\frac{dV_{DM}}{dt} \quad (14)$$

$$\begin{aligned} V_{DM}(r) &= -\int_0^r F_{DM}(r) dr = -\int_0^r \frac{2GM\pi\eta R_{S2}^2}{\sqrt{r}} dr \\ &= -4\pi GM\eta R_{S2}^2\sqrt{r} \end{aligned} \quad (15)$$

Rewriting equation 15 with new variables will be useful in order to have a more simplified Lagrangian. Writing the Lagrangian with the correction is the first step of utilizing the Effective Theories¹:

¹ This section is written by utilizing the methods that were used in [8].

$$V_{DM} = -\frac{\alpha}{r} \left(\frac{R_{efe}}{r} \right)^{-3/2}, \text{ where } \alpha = GMm \text{ and} \quad (16)$$

$$R_{efe} = \left(\frac{4\pi\eta R_{S2}^2}{m} \right)^{-2/3}$$

where m is the mass of S2. The Lagrangian with a \sqrt{r} correction becomes²

$$L_{efe} = \frac{1}{2}mv^2 + \frac{\alpha}{r} \left(1 + \left(\frac{R_{efe}}{r} \right)^{-3/2} \right) \quad (17)$$

Lagrange equations for this correction are

$$\text{Radial: } m \left(\frac{d^2r}{dt^2} - r \left(\frac{d\phi}{dt} \right)^2 \right) = -\frac{\alpha}{r^2} \left(1 - \frac{1}{2} \left(\frac{R_{efe}}{r} \right)^{-3/2} \right) \quad (18)$$

$$\text{Angular: } m \left(2 \frac{dr}{dt} \frac{d\phi}{dt} + r \frac{d^2\phi}{dt^2} \right) = 0 \quad (19)$$

The angular equation can be rewritten as

$$\frac{d}{dt} \left(mr^2 \frac{d\phi}{dt} \right) = 0 \quad (20)$$

Equation 20 indicates that the angular momentum l is conserved, where $l = mr^2 \frac{d\phi}{dt}$. After rewriting $u = 1/r$, changing all time derivatives from $\frac{d}{dt} \rightarrow \frac{d\phi}{dt} \frac{d}{d\phi}$ and replacing $\frac{d\phi}{dt}$ with $\frac{l}{mr^2}$, the radial equation becomes

$$\frac{d^2u}{d\phi^2} + u = \frac{\alpha m}{l^2} \left(1 - \frac{1}{2} R_{efe}^{-3/2} u^{-3/2} \right) \quad (21)$$

To find a solution for u , techniques of the perturbation theory must be employed. The last term in equation 21 is treated as a small perturbation and only the

$$\frac{d^2u}{d\phi^2} + u = \frac{\alpha m}{l^2} \quad (22)$$

part of the equation is solved, which is the standard Newtonian orbit solution:

$$u_N(\phi) = \frac{1}{\rho} (1 + e \cos \phi) \quad (23)$$

² The negative power of the R_{efe}/r suggests that when the semi-major axis of S2 exceeds the value of R_{efe} , the force that is caused by the dark matter radially flowing into the Sgr A* black hole becomes more significant than the force of gravity.

where $\rho = \frac{l^2}{\alpha m}$ (not to be confused with density) and $e = u_0$ (orbit's eccentricity). The perturbation theory states that $u \rightarrow u_N + \delta_{DM}u$ and the equation for $\delta_{DM}u$ is

$$\frac{d^2 \delta_{DM}u}{d\phi^2} + \delta_{DM}u = \frac{\alpha m}{l^2} \left(-\frac{1}{2} R_{efe}^{-3/2} u_N(\phi)^{-3/2} \right) \quad (24)$$

$$= -\frac{1}{2\rho} R_{efe}^{-3/2} u_N(\phi)^{-3/2} = -\frac{\rho^{1/2}}{2} R_{efe}^{-3/2} (1 + e \cos \phi)^{-3/2} \quad (25)$$

The rest of the steps for obtaining an equation that gives the perihelion shift of S2 per orbit were computed using Mathematica; the code can be seen in the Appendices section. However, the steps taken in the code will also be given here in order to give the reader a better understanding. First, a solution to differential equation 25 for $\delta_{DM}u$ is found:

$$\delta_{DM}u(\phi) = u_\delta \quad (26)$$

This solution should be added to u_N in order to obtain $u(\phi)$:

$$u(\phi) = u_N + u_\delta = \frac{1}{\rho} (1 + e \cos \phi) + u_\delta \quad (27)$$

During the perihelions, $u(\phi)$ reaches its maximum, thus the perihelions of the orbit can be obtained by setting $\frac{du(\phi)}{d\phi}$ equal to zero since the derivative of a function becomes zero when it is minimum or maximum. Notice that $\frac{du(\phi)}{d\phi}$ does not become 0 for $\phi = 2\pi$, but it becomes zero for some value $\phi = 2 + \delta_{DM}$. Setting $\frac{du(\phi)}{d\phi}$ to 0 after plugging in $\phi = 2 + \delta_{DM}$, and solving for δ_{DM} would give us our final equation. Note that since δ_{DM} is small and B is also small, the $2.03332B\delta_{DM}$ term becomes too small,³ thus it can be neglected (these terms are in the Appendices section):

$$\delta_{DM} = \frac{21.3242B}{-0.884649A}, \text{ where } B = -\frac{\rho^{1/2}}{2} R_{efe}^{-3/2} \text{ and } A = \frac{1}{\rho} \quad (28)$$

This equation is only valid for S2 since it is obtained by taking $e = 0.884649$, which is the eccentricity of S2.

³ This term can become important for very small e values; however, since we do not work with very small e values, we can neglect this term.

3.1 Writing ρ in terms of Eccentricity and Semi-major Axis

S2 is in an elliptical orbit, which means that its eccentricity is between 0 and 1, $0 < e < 1$. For elliptical orbits and circular orbits

$$r_{min} = \frac{\rho}{1+e} \text{ and } r_{max} = \frac{\rho}{1-e} \quad (29)$$

which results in the semi-major axis (a) becoming

$$a = \frac{r_{min} + r_{max}}{2} = \frac{\rho}{1-e^2} \quad (30)$$

This implies that ρ can be written as

$$\rho = a(1 - e^2) \quad (31)$$

This is more useful since the data for a and e are more precise than the data for mass and angular momentum.

4. Calculations

Table 1. *Data of Stars*

Star Name	Semi-major Axis a (au)	Eccentricity e	Period P (years)	Mass (m) M_{\odot}	Radius R_{S2} (au)
S2	1031 [9]	0.884649 [9]	16.0 [9]	15 [10]	0.03258 [11]
S38	1160 [12]	0.8201 [12]	19.2 [12]	ND	ND
S55	900 [12]	0.7209 [12]	12.8 [12]	ND	ND
S62	746 [13]	0.9760 [13]	9.9 [13]	6.1 [13]	0.02414
S4711	619 [13]	0.768 [13]	7.6 [13]	2.2 [13]	0.01718
S4714	841 [13]	0.985 [13]	12.0 [13]	2.0 [13]	0.01664
S4715	1187 [13]	0.247 [13]	20.2 [13]	2.8 [13]	0.01862

In Table 1, the source of the data is indicated in each of the entries. ND indicates that no data was found. Additionally, we were only able to find data for the radius of S2, and the radii of the 10 other stars were found by accepting that all of the S-stars have approximately the same density.

The collected data from articles and websites for the semi-major axis, a , was in the unit of arcseconds. In order to change the units into astronomical units, a set of operations was performed. Arcseconds as a unit of length represent the angle between two objects viewed from the Earth. When the trigonometry is performed, it is clear that the semi-major axis of S2 can be written as

$$a = 2R_{GC} \sin \theta/2 \quad (32)$$

where R_{GC} is the distance to the galactic center from Earth (8.2467 kpc or 1.701109 au) and θ is the semi-major axis in the unit of arcseconds. Each star is accepted to be equidistant from the Earth and the distance from Earth to the stars is accepted to be equal to R_{GC} . Due to small-angle approximation, equation 32 can be rewritten as

$$a = R_{GC} \quad (33)$$

where R_{GC} should be in the preferred units (au) and θ should be translated into radians by being multiplied with $\pi/648000$.

4.1 Schwarzschild Radius

Table 2. Data of Sgr A*

Mass m ($M_{\odot} \cdot 10^6$)	Schwarzschild Radius R_{SC} (au)
4.261 [9]	0.084

In Table 2, the source of the data is indicated in each of the entries.

The Schwarzschild Radius is calculated via the following formula:

$$R_{SC} = \frac{2GM}{c^2} \quad (34)$$

where c is the speed of light.

4.2 Constraints on η

Most recent observations made by the GRAVITY suggest that the perihelion shift of S2 can be expressed by the following equation [9]:

$$\delta_{Total} = f_{SP} \cdot \delta_{GR} \quad (35)$$

The perihelion shift predicted by General Relativity can be found using the following equation:

$$\delta_{GR} = \frac{3\pi R_{SC}}{a(1 - e^2)} \quad (36)$$

For S2, the relativistic precession becomes 12.14 arcminutes per orbit. The data suggests that the constant f_{SP} is 1.1, with an uncertainty of ± 0.21 . Thus, the maximum precession of S2 within the error bars is about 15.9 arcminutes per orbit. Dark matter wind is predicted to also cause S2 to precess, which can be a maximum of 3.76 arcminutes per orbit of S2. This enables us to put constraints on the constant η . In order to find the maximum value of η , first B should be found

from equation 28. The constant A in the same equation can be found from the observed data:

$$A = \frac{1}{\rho} = \frac{1}{a(1 - e^2)} \approx 0.00446 \text{ au}^{-1} \quad (37)$$

We can find B by rewriting equation 28 and plugging in A :

$$B = \frac{-0.884649A\delta_{DM}}{21.3242} \approx -2.0244 \cdot 10^{-7} \text{ au}^{-1} \quad (38)$$

Now, all we need to do is to rewrite equation 28, and plug in the data and calculated constants to find R_{efe} ⁴:

$$R_{efe} = \left(-\frac{\rho^{1/2}}{2B} \right)^{2/3} = \left(-\frac{(a(1 - e^2))^{1/2}}{2B} \right)^{2/3} \approx 110989.6 \text{ au} \quad (39)$$

Finally, in order to get η , equation 16 should be rewritten, and the value obtained for R_{efe} and the gathered data should be plugged in:

$$\eta = \frac{R_{efe}^{-3/2} m}{4\pi R_{S2}^2}, \quad \eta_{max} \approx 3.041 \cdot 10^{-5} M_{\odot}/\text{au}^{9/2} \quad (40)$$

Note that this equation is only valid for the dark matter density around the Sgr A* central black hole, and can not be used to estimate the density of dark matter around stars such as our Sun, which are far away from the galactic center. Equation 40 suggests that the dark matter density near S2 is less than $9.765 \cdot 10^{-4} M_{\odot}/\text{au}^3$, or, in more familiar units, $5.8 \cdot 10^{-10} g/cm^3$. When it is compared to the upper limit for the local dark matter density, which is about $1.4 \cdot 10^{-19} g/cm^3$, the local dark matter density is about $4.1 \cdot 10^9$ times smaller than the maximum possible dark matter density around the S2 [14]. This is an expected result since the Earth is much further away than the galactic center compared to S2 and dark matter density is expected to be higher near the galactic center.

4.3 Predictions for other S-stars

There are dozens of S-stars orbiting Sgr A*; however, only a few of them are shown in Table 1. The main reason for this is the orbital period. Most of the S-stars have orbital periods of over one hundred years, and only stars with orbital periods of less than 25 years are included in Table 1. These stars with relatively short orbital

⁴ Footnote 2 suggested that if the semi-major axis of S2 was larger than R_{efe} , the dark matter force was more important than gravity. However, $R_{efe} \approx 108a$, which suggests that the dark matter force is as significant as gravity for stars that are about 108 times further from Sgr A* than S2. Equation 39 also implies that for S2, the potential energy created by the dark matter force is about 1117 times smaller than the Newtonian gravitational potential energy.

periods are selected since they might be the ones on which new measurements could be made in the near future.

The constraints that we have put on the dark matter density at the center of our galaxy enable us to put upper limits on the δ_{DM} and f_{SP} values of other S-stars. In order to make these predictions, equation 28 should be rewritten for each star. Equation 28 is only valid for S2 since in order to obtain that equation the eccentricity of S2 was plugged in for e . This means that, in order to find the equations that will give the δ_{DM} of other stars, all we need to do is plug their eccentricity values into the code. After obtaining the equations, all we need to do is plug in the data. The results for other S-stars are shown in Table 3⁵.

Table 3. Predictions for Other Stars Based on the Data from S2

Star Name	S62	S4711	S4714	S4715
δ_{GR}	76.95	10.72	108.72	2.44
δ_{DM-max}	1.53	4.40	2.12	15.28
f_{SP-max}	1.02	1.41	1.02	7.26

It is clear that among these four stars S4715 is the best candidate for testing this effect. If the predicted interaction between dark matter and stars is correct, the data gathered from instruments such as GRAVITY should yield a result for the perihelion shift of S4715 over seven times larger than the prediction. An error margin of this magnitude would clearly indicate that General Relativity by itself is not enough to make valid predictions in some cases. The only downside of gathering data for S4715 is that it is the star with the largest orbital period among these four stars, thus the data gathering would take quite some time.

5. Discussion

In this paper, we calculated an upper limit for the density of dark matter at the center of our galaxy from the measurements made by the GRAVITY for the perihelion shift of S2. Even though the data have enabled us to approximate an upper limit, the uncertainties were significant and S2 was not the best star to test the predicted effect. Even though the constant f_{SP-max} had a substantial error bar, it was not the only significant source of uncertainty. The data for the mass of S2 fluctuated from website to website and the data for S2's radius was approximated significantly. There were other approximations made, such as the distance from Earth to S2, but these approximations had very minimal effects on the final result. The future measurements for mass and radius of S2 and other stars would enable us to put a more reliable upper limit on the dark matter density at the center of our universe.

As mentioned before, S2 was not the ideal star to test this effect. As can be seen in Table 3, S4715 yields the greatest f_{SP-max} value among other S-stars with orbital periods less than 25 years. Stars with high f_{SP-max} values should be studied because a high value for f_{SP} indicates that there must be another reason for the star to experience perihelion precession. Thus, studying S4715 in the

⁵ S38 and S55 could not be tested for this effect because the data for their mass and radius could not be found.

future and finding the actual f_{SP} value for S4715 would enable us to test our hypothesis much better.

There are several reasons for S4715 having a high hypothesized f_{SP-max} value. High f_{SP} occurs when the effect of general relativity is low and the effect of dark matter is high. δ_{DM} is inversely proportional to both $(1 - e^2)$ and a . Thus, the effects of general relativity are most visible in the orbits with high eccentricity and low semi-major axis. It is the exact opposite case for δ_{DM} . From equation 28, it is clear that δ_{DM} is proportional to $a^{3/2}$ however, the relationship between eccentricity and δ_{DM} is a bit more complex. Equation 28 is obtained by plugging in a number for eccentricity, thus we can not directly conclude that $(1 - e^2)^{3/2}$ and δ_{DM} are proportional. The numerical constants in equation 28 seem to change in a way that increases δ_{DM} as the eccentricity increases. However, the decrease that comes from $(1 - e^2)^{3/2}$ is more dominant.

To conclude, even though the nature of the relationship is a bit ambiguous, δ_{DM} increases as eccentricity decreases⁶. Returning to the initial point, in orbits where the effect of dark matter is more significant, the effect of General Relativity is less significant compared to other orbits. The listed reasons make S4715 the best candidate to test our theory since it has a relatively big semi-major axis and a relatively small eccentricity compared to other S-stars with short orbital periods. Another naive reason for S4715 being a good choice for testing this effect is that its mass is not too big. This last reason is only valid due to an approximation we made. We accepted the density of all S-stars to be the same, thus the R^2/m factor in equation 28 (which is not directly in the equation, but this factor is in R_{efe}) becomes larger as mass decreases. Thus, this final reason for selecting S4715 is not as important as the others before it, but it would still be interesting to look at the densities of stars in the future and find a relationship between the density of a star and δ_{DM} .

In the previous paragraph, we pointed out two variables that affect δ_{DM} . In this paragraph, we will try to explain the reasons behind these relationships. As stated before, δ_{DM} increases as the semi-major axis of a star increases. The main reason for this could be the distribution of dark matter around Sgr A*. We concluded that the power of r in equation 10 should be $\frac{1}{2}$, resulting in dark matter density being dependent on \sqrt{r} . Thus, in orbits with larger semi-major axes, the density of dark matter around stars would be more than the density of dark matter around stars with smaller semi-major axes. The increased density means more dark matter interacts with the star and more force is applied to the star. A larger force indicates a larger perihelion shift, thus this relationship between δ_{DM} and the semi-major axis is expected. Again as stated before, δ_{DM} increases as the eccentricity of a star's orbit decreases. A reason for this might be the change in the time that a star spends in dark matter-dense regions. In highly eccentric orbits, stars spend more time further away from the black hole than closer to it due to the conservation of angular momentum. However, this does not mean that stars

⁶ This relationship between e and δ_{DM} is not valid for extremely small eccentricities. δ_{DM} increases as e decreases; however, for very small numbers, the term neglected while obtaining equation 28 becomes significant, and the term with A starts to become insignificant. This is not important since the equation is obtained perturbatively.

spend more time in dark matter-dense regions. Since dark matter density increases as a function of \sqrt{r} , it has a strange distribution around Sgr A*; the increase in density decreases as r increases. This might result in highly eccentric orbits interacting with less dark matter than stars with less eccentric orbits.

Additionally, the described interaction between dark matter and stars orbiting Sgr A* is only one possible scenario among many. Dark matter particles might also be orbiting Sgr A* like other stars, and in this case, the effect of dark matter on the orbital dynamics of the stars would be much more complex. Dark matter would not only collide with the stars, but it might also attract them gravitationally. The collisions that would take place would not be as simple as the one that is speculated in this paper; the transfer of momentum would not only be in the r -direction, but could also be in the ϕ -direction. The second scenario might result in the orbital velocity of the star change; and if the change is significant, its effect might be observable in the form of an increased semi-major axis, etc.

References

- [1] Dark matter, URL <https://home.cern/science/physics/dark-matter>.
- [2] A. Riess, dark matter, URL <https://www.britannica.com/science/dark-matter>.
- [3] Dark matter, dark energy, URL <https://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy>.
- [4] G. Bertone and D. Hooper, *Reviews of Modern Physics* 90 (2018), ISSN 1539-0756, URL <http://dx.doi.org/10.1103/RevModPhys.90.045002>.
- [5] T. Talks, On light and dark matter: David Kaplan at Tedxjohnshopkinsuniversity (2014), URL <https://www.youtube.com/watch?v=aUohYDSJ24g>.
- [6] L. Bergström, *Reports on Progress in Physics* 63, 793–841 (2000), ISSN 1361-6633, URL <http://dx.doi.org/10.1088/0034-4885/63/5/2r3>.
- [7] D. J. Mortlock and E. L. Turner, *Monthly Notices of the Royal Astronomical Society* 327, 557 (2001), ISSN 0035-8711, <https://academic.oup.com/mnras/article-pdf/327/2/557/3514307/327-2-557.pdf>, URL <https://doi.org/10.1046/j.1365-8711.2001.04774.x>.
- [8] J. D. Wells, When effective theories predict: the inevitability of mercury's anomalous perihelion precession (2011), 1106.1568.
- [9] R. Abuter, A. Amorim, M. Bauböck, J. P. Berger, H. Bonnet, W. Brandner, V. Cardoso, Y. Clénet, P. T. de Zeeuw, and et al., *Astronomy Astrophysics* 636, L5 (2020), ISSN 1432-0746, URL <http://dx.doi.org/10.1051/0004-6361/202037813>.
- [10] Hint of relativity effects in stars orbiting supermassive black hole at centre of galaxy (2017), URL <https://www.eso.org/public/announcements/ann17051/>.
- [11] S2 and central black hole, URL <http://www.solstation.com/xobjects/s2.htm>.
- [12] S. Gillessen, P. M. Plewa, F. Eisenhauer, R. Sari, I. Waisberg, M. Habibi, O. Pfuhl, E. George, J. Dexter, S. v. Fellenberg, et al., *The Astrophysical Journal* 837, 30 (2017), ISSN 1538-4357, URL <http://dx.doi.org/10.3847/1538-4357/aa5c41>.

[13] F. Peiβker, A. Eckart, M. Zajačec, B. Ali, and M. Parsa, *The Astrophysical Journal* 899, 50 (2020), ISSN 1538-4357, URL <http://dx.doi.org/10.3847/1538-4357/ab9c1c>.

[14] N. P. Pitjev and E. V. Pitjeva, *Astronomy Letters* 39, 141–149 (2013), ISSN 1562-6873, URL <http://dx.doi.org/10.1134/S1063773713020060>. 17

Appendix

```

usol = DSolve [D[D[u[φ], φ], φ] + u[φ] == B (1 + e Cos[φ])^(-3/2), u[φ], φ]
Out[1]= {{u[φ] → c1 Cos[φ] + c2 Sin[φ] -
      (2 B (-Cos[φ] + e2 Cos[φ] - √(1 + e Cos[φ]) / (1 + e) EllipticE[φ/2, 2 e / (1 + e)] Sin[φ] - e √(1 + e Cos[φ]) / (1 + e) EllipticE[φ/2, 2 e / (1 + e)] Sin[φ] + √(1 + e Cos[φ]) / (1 + e) EllipticF[φ/2, 2 e / (1 + e)] Sin[φ] - e2 √(1 + e Cos[φ]) / (1 + e) EllipticF[φ/2, 2 e / (1 + e)] Sin[φ] + e Sin[φ]2) / (e (-1 + e2) √(1 + e Cos[φ]))}}

usolfull = usol /. {C[1] → 0, C[2] → 0}
Out[2]= {{u[φ] → -((2 B (-Cos[φ] + e2 Cos[φ] - √(1 + e Cos[φ]) / (1 + e) EllipticE[φ/2, 2 e / (1 + e)] Sin[φ] - e √(1 + e Cos[φ]) / (1 + e) EllipticE[φ/2, 2 e / (1 + e)] Sin[φ] + √(1 + e Cos[φ]) / (1 + e) EllipticF[φ/2, 2 e / (1 + e)] Sin[φ] - e2 √(1 + e Cos[φ]) / (1 + e) EllipticF[φ/2, 2 e / (1 + e)] Sin[φ] + e Sin[φ]2) / (e (-1 + e2) √(1 + e Cos[φ]))}}

In[3]= usolfull[φ] = A (1 + e Cos[φ]) + u[φ]
Out[3]= A (1 + e Cos[φ]) + u[φ]

In[4]= Series[D((usolfull[φ] /. usolfull), φ) /. {φ → 2 Pi + δ}, {δ, 0, 1}]
Out[4]= {
      4 B (EllipticE[2 e / (1 + e)] - EllipticK[2 e / (1 + e)] + e EllipticK[2 e / (1 + e)]) (2 B + 3 B e - A e2 (1 + e)3/2) δ
      / ((-1 + e) e √(1 + e)) + (2 B + 3 B e - A e2 (1 + e)3/2) δ / (e (1 + e)3/2) + O[δ]2
}

Solution for S2
In[5]= Series[D((usolfull[φ] /. usolfull), φ) /. {φ → 2 Pi + δ, e → 0.884649}, {δ, 0, 1}]
Out[5]= {-21.3242 B + (-0.884649 A + 2.03332 B) δ + O[δ]2}

Solution for S62
In[10]= Series[D((usolfull[φ] /. usolfull), φ) /. {φ → 2 Pi + δ, e → 0.9760}, {δ, 0, 1}]
Out[10]= {-113.275 B + (-0.976 A + 1.81778 B) δ + O[δ]2}

Solution for S4711

```

Figure 3. Mathematica code first page

```
In[12]:= Series[D[ufull[phi] /. usolfull], phi] /. {phi -> 2 Pi + delta, e -> 0.768}, {delta, 0, 1}]
Out[12]:= {-9.46761 B + (-0.768 A + 2.3839 B) delta + O[delta^2]}

Solution for S4714

In[13]:= Series[D[ufull[phi] /. usolfull], phi] /. {phi -> 2 Pi + delta, e -> 0.985}, {delta, 0, 1}]
Out[13]:= {-183.526 B + (-0.985 A + 1.79873 B) delta + O[delta^2]}

Solution for S4715

In[11]:= Series[D[ufull[phi] /. usolfull], phi] /. {phi -> 2 Pi + delta, e -> 0.247}, {delta, 0, 1}]
Out[11]:= {-1.24684 B + (-0.247 A + 7.96916 B) delta + O[delta^2]}
```

Figure 4. Mathematica code second page



The Consequences of Urbanization on Access to Housing and Education in Turkey: Struggles and Coping Strategies of Syrian Refugees

Ezgi Erzincan

Author Background: *Ezgi Erzincan grew up in Turkey and currently attends Lycée Saint Joseph in Istanbul, in Turkey. Her Pioneer research concentration was in the field of sociology and titled “Globalization and International Migration.”*

Abstract

This interdisciplinary research project aims to demonstrate how Syrian immigrants and refugees living in Turkey experience problems accessing housing and education because of intra-rural and external migration’s effects on urbanization and to spotlight select initiatives of the government and NGOs.

A spike in the urban population has severely affected housing availability to the point that almost all developed land and apartments have been occupied and unauthorized slums have sprung up. Syrian immigrants in Turkey move to city centers from unsanitary camps where they end up sharing small accommodations because of exorbitant rental costs and all the while confronting ever-increasing housing prices and discrimination because of their Syrian identity. Alongside housing, access to education plays a key role in the integration of migrants. Deterring factors such as distance to schools and cost of transportation to commute to a language course center prevent many Syrian immigrants from pursuing (higher) education upon arrival to Turkey. Children who live in slums are reported to quit school to earn money through hard labor and thus run the danger of living without the ability to read or write.

As a Turkish woman with a strong belief in the need to protect human rights and who has been a witness to this migration process from the beginning, I will explore to what extent immigrants’ lives are affected by the urbanization process with a focus on housing and access to education. Specifically, I seek to answer the following questions, “How is urbanization, caused by external and intra-Turkish

migration, affecting the accessibility of housing prospects and educational opportunities to Syrian immigrants residing in Turkey with an emphasis on large cities like Istanbul?” and, “What resources exist to assist this vulnerable community?” For this research project, I conducted interviews with the Senior Monitoring Evaluation Accountability Officer at MSYD-ASRA and the Project Officer at “The Research Centre on Asylum and Migration (İGAM)” to gain a better understanding of the struggles Syrian refugees face in regards to housing and access to education, as well as what steps are being taken to improve these conditions. I argue that Syrian refugees in Turkey are affected negatively by urbanization when it comes to finding housing opportunities because of lack of land in cities, extortionate rents, and unsafe housing structures as well as the fact that that Syrian refugee children’s access to education has been disrupted as a consequence of these housing challenges.

I conclude that the challenges posed by urbanization — primarily limited housing stock in city centers and the emergence of slums in under-resourced neighborhoods — has necessitated Syrian refugees to embrace strategies that are in the long-term harmful for their families, for example, multiple families stay in one run-down home to afford rent, send their children to work for additional family income, stop girls from going to school because of safety or cultural concerns, and have more children to receive more government aid. At the same time, in an effort to mitigate the challenges faced by refugees, NGOs and the Turkish government have set up comprehensive nationwide support programs to ensure that refugee children stay at school and that their families can afford basic necessities more easily.

1. Introduction

As of May 26th, 2021, 3,672,646 Syrian refugees have been reported to reside in Turkey (Mülteciler Derneği 2021). The period in which most immigration to Turkey took place was during the Syrian Civil War years (Bimay and Kaypak 2016, 91-92). This influx of millions of people has caused significant challenges for the Turkish government as it needed to provide, in a short period of time, dignified accommodation, health care, and education to a population traumatized by displacement, adding to its already significant economic and political struggles (Salihi 2016, 159).

In particular intra-Turkish migration over the past decades has caused cultural, social, economic, and political changes in Turkey. Notably, one of the biggest consequences of this internal migration was increased urbanization, which put pressure on already scarce housing stock and limited access to shelter and education for Syrian immigrants.

As of July 25th, 2019, 87,464 Syrians (only 2.4%) are reported to live in refugee camps while 3,552,004 (97.6%) official immigrants are reported to live in urban environments and city centers (Hamsici 2019). Because the camps are overcrowded, Syrian immigrants preferred staying at rental properties in the cities. Reports (2020) showed groups of 10-12 people put together in 10-15 square meter spaces without sufficient infrastructure. They also stated that immigrants living in camps could not get assigned a tent and were forced to sleep on the floor of such camps. (Pişkin 2020). Another study (2014) showed that in Gaziantep, many Syrians lived in sheds and slums (Euronews 2014). Another effect of urbanization on housing is reported to be an increase in rental prices. Syrians are preferred by landlords to rent apartments because they can afford to pay higher rental fees. As the demand for housing properties and basic supplies increased, so did their prices (ORSAM & TESEV 2015, 17-19). A 2019 research paper by Doğuş Şimşek demonstrated that Syrians struggled to pay high rental fees, and many families stayed together just to afford apartments with poor conditions (Şimşek 2018, 379).

The effects of urbanization have also affected the accessibility of education for Syrian immigrants and their children. There is limited knowledge and research on the education level of the Syrian immigrant population in Turkey (Erdoğan 2020). In addition, children who live in slums reportedly quit school to earn money through hard labor and have a higher propensity towards illiteracy and early marriage.

By bringing the unexplored struggles regarding housing and access to education of Syrian immigrants living in Turkey into light, I will also seek to uncover what steps are being taken to improve housing opportunities and the accessibility of education for Syrian immigrants and show to what extent these immigrants' lives are affected due to urbanization altering housing and access to education.

Following the introduction, in section 2 of this paper, I will provide background on the process of urbanization in Turkey, accelerated by intra-Turkish rural to urban migration, and the Syrian refugee crisis. Additionally, I will offer a snapshot of the laws and policies surrounding the treatment and rights of Syrian refugees who reside in Turkey. In part 3, I will present information on the housing market and access to education as it relates to Syrian refugees. In the following section, I will describe and evaluate the services self-help organizations are offering in terms of supporting Syrian refugees. I will conclude with final reflections and recommendations on how to overcome the housing and education problems of Syrian refugees.

2. The Effects of Intra-Turkish Migration and the Syrian Refugee Crisis on the Urbanization Process in Turkey

Urbanization is the process of the population shifting from the countryside and villages to cities. In a study done on the Yeşilyurt district, the terms “repulsiveness of the rural” and “urban attractiveness” were clearly defined as the main drivers for urbanization. “Attractive” factors of the city such as

better employment opportunities, high-quality education, easy transportation, and availability of better health care are overlaid to the “repulsive” factors of the countryside such as insufficiency of income, poverty, fragmentation of lands by inheritance, marketing problems of products, and non-agricultural work (Aşkın, Yayar, and Oktay 2013, 231).

Urbanization in Turkey was triggered by the industrial revolution and became fully realized after World War II. The process was a rapid one (Yıldırım 2004, 24-25). In the first years after the establishment of the Republic in 1923, a large portion of the population of Turkey lived in villages. This ratio has shown development in favor of cities over the years. While the rate of population migrating between settlements was 9.34% in 1980, this rate was 8.67% in 1985, 10.81% in 1990, and 11.02% in 2000. The proportion of people living in provinces and districts in Turkey increased to 93% in 2020 (TÜİK 2021). This change in population ratios had consequences in many different fields. As the population density increased, there was a housing shortage in the cities and slums, and squatting increased. In addition, existing infrastructure, health providers, and education became insufficient (Terzi and Koçak 2012, 166-169). When the data of TÜİK (Turkish Statistical Institute) are analyzed and the phenomenon of urbanization is considered with its demographic dimension, it is seen that the number of individuals living in cities in Turkey is increasing with each passing year (Zengin 2018, 85).



Figure 1. *Distribution by the population of the size of the cities with over 50.000 people in Turkey (1950)*

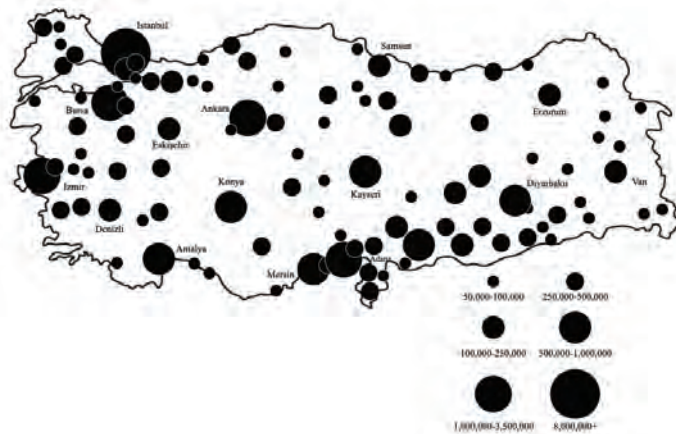


Figure 2. Distribution by the population of the size of the cities with over 50.000 people in Turkey (2000) (Işık 2005, 69-71)

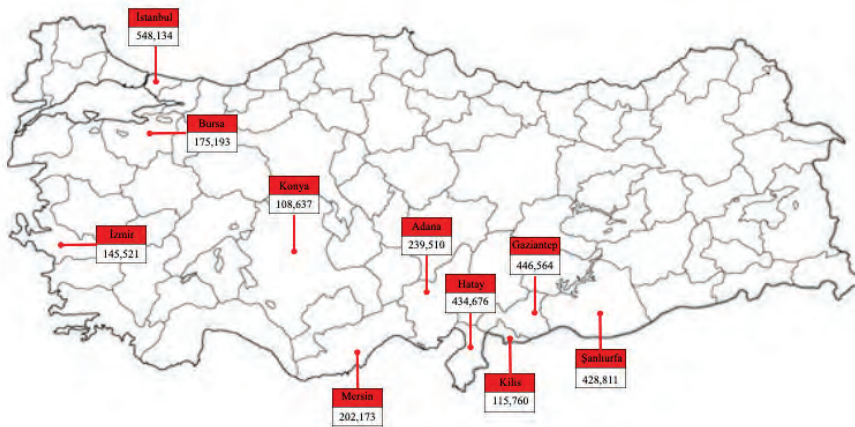


Figure 3. Distribution of Syrians Under Temporary Protection by Top Ten Cities (Erdoğan 2019, 6)

Alongside intra-Turkish rural to urban migration, the increase in the Syrian refugee population in Turkey contributed to the process of urbanization in the last decade. The Syrian Civil War resulted in many Syrians and their families fleeing the country to seek asylum in neighboring countries, including Turkey. Turkey opened its first refugee camp for Syrian refugees in May 2011. With its open-door policy, Turkey welcomed any Syrian refugees seeking protection and housing. However, after a series of terrorist attacks committed by Daesh in 2015 and 2016 which endangered the lives of many citizens and visitors, the Turkish government put in place certain restrictions such as closing its borders as of 2016. Due to these constraints, the number of refugees started to decline. According to the U.N. High Commissioner for Refugees (UNHCR), as of February 7, 2019, there

were a total of 3,644,342 registered Syrian refugees in Turkey. Currently, 15% of Syria's pre-Syrian civil war population now lives in Turkey (Makovsky 2019). The August 2020 data of the Ministry of Interior, General Directorate of Migration Management, reported that 98% of Syrian refugees registered in Turkey currently reside in cities (Özer and Beyazıt 2020, 548), with the majority living in Istanbul, Gaziantep, Hatay, Şanlıurfa, and Adana. The Syrian refugees, who fled from the war to save their lives, generally came to the city centers, except for certain provinces, due to their migration process happening too suddenly. They were drawn to cities because they believed there would be job opportunities there. The only exceptions to this were the refugees who entered the agricultural sector as cheap workers or the people in the border provinces taking their relatives to their villages. Additionally, there were certain reasons for Syrians who preferred living in border provinces. These reasons had a direct connection to the fact that the countries near the border had large populations speaking Arabic and that the lack of knowing Turkish did not pose as big of a problem there compared to other regions (Eren and Aydın 2016, 35-36). Considering these data which have direct effects on cities, an increase in Syrian refugees and intra-Turkish rural to urban migration have had consequences on urbanization policies.

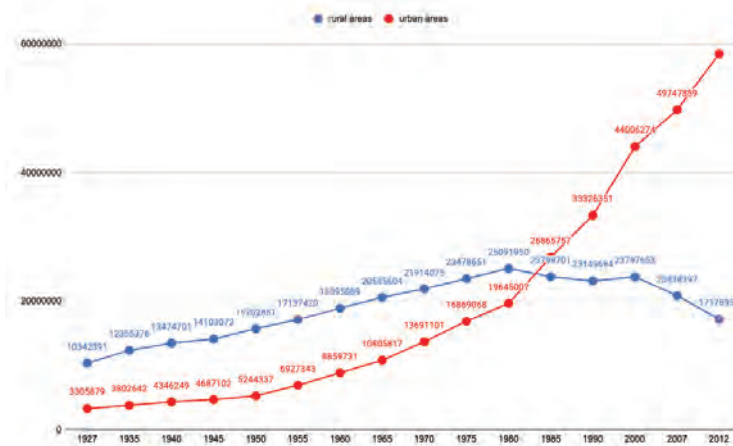


Figure 4. *Change of Rural and Urban Population in Turkey by Years (Yılmaz, n.d., 165)*

3. Laws and Policies Surrounding the Treatment and Rights of Syrian Refugees

Between 1934 and 2006, the Law on Settlement limited settlement in Turkey into two statuses: facilitating the migration and integration of people of “Turkish origin and culture” as migrants or refugees and preventing entry of people who did not meet this criterion as migrants or refugees (İçduygu and Aksel 2013, 171-172). Even after the new Law on Settlement was implement-

ed and subjected to change, it remained the core document of the settlement policy of the Republic of Turkey (İnan 2016, 24).

The Law on Work Permits of Foreigners set rules as to how foreigners could work with employers to apply for a job. The law states that for a refugee or a foreigner to be eligible for a job, at least five Turkish citizens must also be working at the corporation.

Another important law, the Temporary Protection Regulation, was issued on October 22, 2014. This regulation allowed temporary protection towards refugees who were in an emergency situation and who could not go back to their countries. (Library of Congress 2020). The rights granted through Temporary Protection range from social benefits to educational services. Refugee women who have been subjected to violence are permitted to reside in women's shelters upon filing a complaint. In addition, refugee children can enroll in Turkish universities after fulfilling certain requirements. Syrians under Temporary Protection can also participate in free courses offered by Public Education Centers, including language development courses, skill development courses, hobby courses, and vocational courses.

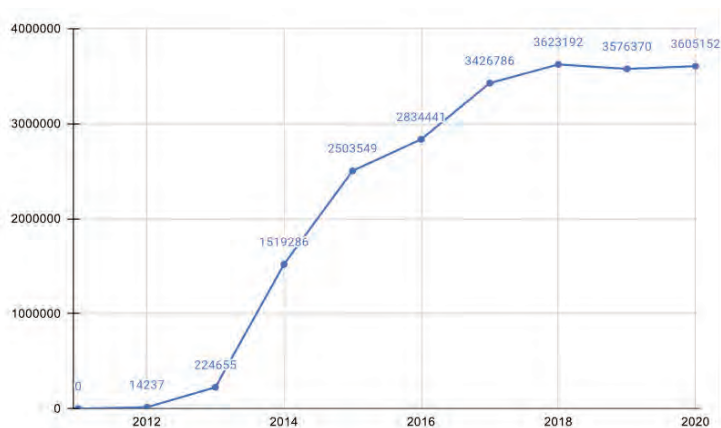


Figure 5. Number of Syrians Under Temporary Protection in Turkey (Çalışkan 2020, 1195)

The laws of the Republic of Turkey state that all children, regardless of their status as a citizen or a foreigner, are allowed to access primary and secondary education opportunities without pay. Children who are in need of additional educational support are also given Turkish language lessons outside school hours (UNHCR 2017, 6-29)

While the laws and rights given to Syrian refugees are nowhere near as generous as those open to citizens, the general public has many misconceptions that arise from a lack of education on the subject. Some common beliefs and misconceptions about Syrian refugees include that Syrians are not obligated to pay taxes, have the ability to vote without being citizens, and take the same university entrance exam as Turkish citizens. As a result, many mass violence incidents and discrimination from Turkish citizens towards Syrian refugees happened.

Stephan (2000) shows how prejudices form and can be overcome, and introduced the Integrated Threat Theory that highlights how fear and the perception of threat play roles in intergroup relations (Stephan 2000, 24). This study shows that perception of threat from a group is the root for negative treatments towards the other group, regardless of whether the threat is factual. The study can help understand the negative attitudes towards Syrian refugees living in Turkey, who are presumed to “steal” the jobs of citizens of Turkey (Özüdoğru et al. 2018, 121). Due to the general economic crisis in the country, economic factors such as increasing unemployment and competition in the labor market are the culprit for the perception of the threat of Istanbul residents towards Syrian refugees. As Syrians become more integrated into business and companies, Turkish citizens do not feel as much pity for the Syrian refugees (Tokyay 2020).

4. The Housing Situation of Syrian Refugees

The Syrian refugee population is concentrated in large cities and thus puts great pressure on the local real estate market. As a result, housing prices have been rising rapidly. In addition, the competition between the poor resident population and newcomers for cheap housing with insufficient amenities in the slums becomes fiercer (Şentürk 2020, 17).

In research completed on Mersin, a city in the Mediterranean region of Turkey, some of the biggest issues surrounding housing and housing spaces in the city were reported. The housing market which is in high demand due to Syrian refugee numbers multiplying each year, combined with a lack of organizations that can provide consultancy to Syrian refugees regarding housing drove up the rental prices very rapidly, by 20-30%. Syrian refugees do not have non-governmental organizations who can inform and assist them with actions like renting or purchasing a property. Alongside these issues, it is also a widely known problem that landlords take advantage of Syrian tenants by receiving unfair commissions from them (Mersin University and Doğan 2014, 63-64). In addition, the xenophobia of the local residents towards Syrian refugees can sometimes prevent them from being able to rent a certain property.

4.1 Syrian Refugees' Journey from Temporary Accommodation Centers to City Centers

For refugees and asylum seekers, access to suitable housing is a sign of successful integration and newcomers can start a new life by living with people living in the host country (Bakbak 2018, 254). When Syrian refugees arrived in Turkey, they were initially offered residency in temporary accommodation centers as their duration of stay in Turkey was seen as only temporary. With many immigrants permanently settling in cities as a result of uncomfortable living conditions at temporary accommodation centers, there are now very few who still reside in these places. On the subject of the recent situation of refugee camps, Mr. Tecer, Senior Monitoring Evaluation

Accountability Officer at MSYD-ASRA, explained that there is no current overcrowdedness in the camps and that the majority of refugees reside in the city centers.

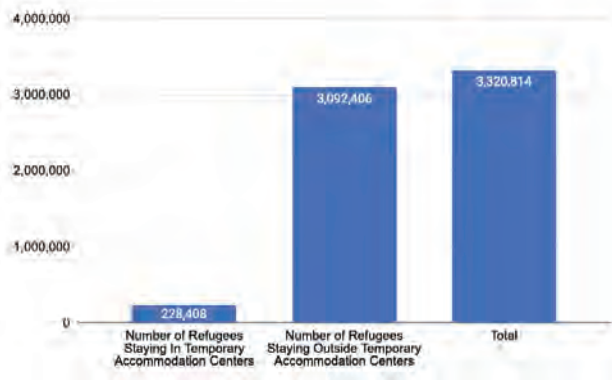


Figure 6. *Syrians Staying in and out of Temporary Accommodation Centers (2017) (Taş and Özcan 2018, 45)*

The proportion of refugees who stay in camps in Turkey is very low. It's almost 2% to 1%. Staying in camps is not a priority for a refugee (in Turkey). These camps also don't have such capacity. Most of them are located in Hatay, Gaziantep, Urfa, Kilis, Kahramanmaraş. The vast majority of them are closed and (the refugees) have become urban immigrants. Only 55,743 people live in the camps. Since the number of refugees living in the city is very high, they come to the city centers and centers where industrialized job opportunities are plenty. In the places they stay, they have large family structures. In these large family structures, there are many orphans and families with orphaned children. Amongst these are also those who stay with their relatives (there are also structures where more than 1 or 2 families stay together). We even encountered families of 20 people in a household and (their houses) didn't have enough rooms. They face many dangers. They may face dangers such as abuse or invasion of their private life. There is no overcrowding in the camps. Currently, staying in camps is not a preferred situation as there are neither new migration flows compared to the past nor immigrants who want to return to these camps. (Tecer 2021) ¹

¹ (Turkish Transcription) "Türkiye'deki kampta kalan mülteci oranı çok az. Neredeyse yüzde 2 yüzde 1 civarında. Kamplarda kalmak bir mülteci için (Türkiye'de)

4.2 Regions and Types of Houses in which Syrian Refugees Find Housing Opportunities

In 2015, a non-governmental organization whose name was not disclosed conducted research on refugees' housing patterns on the southern border of Turkey with 36 thousand Syrian refugees. Their research found that only 33.3% of Syrian refugees live in modern and insulated buildings. The rest of the population lives in damp houses with insufficient insulation and heating problems (43.3%), buildings that are not suitable for permanent shelters, such as tents (17.9%), garages or unfinished construction (4.96%), and public areas such as mosques and parks (Amnesty International 2016, 25). In our conversation, Mr. Tecer listed easier access to job opportunities and more affordable houses with improper structures as the main reasons why Syrian refugees prefer to stay in cities, highlighting the regions in which Syrian refugees tend to group:

We observe Syrian refugees factioning in cities, edges, and peripheries of cities, big cities that have been subjected to urbanism, and regions that we can call suburbs of metropolises. For example, the reason for grouping in the Bağcılar and Esenler regions in Istanbul is the need for cheap housing and easy access to the workforce. These regions come up as places with under-the-counter workshops in terms of textiles, outdated houses with cheaper rents that will be exposed to urban transformation, and houses that are yet to have proper infrastructure. We have seen (*issues*) ranging from sewer problems, humidity inside the house, and houses without roofs, windows, or doors. (Tecer 2021) ²

bir öncelik değil. Öyle bir kapasiteleri de yok. Büyük bir çoğunluğu Hatay, Gaziantep, Urfa, Kilis, Kahramanmaraşta. Bunların büyük bir çoğunluğu kapatıldı ve bunlar şehir göçmenleri haline gelmeye başladı. 55 743 kişi kampta yaşıyor sadece. Şehirde yaşayan mülteci sayısı çok fazla olduğu için şehir merkezlerine ve sanayileşmiş iş olanaklarının fazla olduğu merkezlere geliyorlar. Kaldıkları yerlerde geniş aile yapılarına sahipler. Bu geniş aile yapılarında da birçok yetim, öksüz çocuklu aileler de oluyor. Bunlardan akrabalarıyla birlikte kalanlar da var (1'den 2'den fazla ailenin bir arada kaldığı yapılar da var). Bazen bir hane içerisinde 20 kişilik ailelerle karşılaştığımız da oluyordu ve bunlar yeterli sayıda odaya sahip değiller. Birçok tehlikeyle karşı karşıyalar. Hem istismar, hem özel hayatın doğru düzgün yaşanması gibi tehlikelerle karşı karşıya kalabiliyorlar. Kamplarda aşırı kalabalıklaşma söz konusu değil. Şu anda da geçmişe nazaran ne yeni göç akımı ne de kampa dönmek isteyen göçmenler olmadığı için tercih edilen bir durum değil.”

² (Turkish Transcription) “Kentlerde, kentlerin kenarında ve çeperinde diyebileceğimiz kentsel dönüşüme maruz kalmış büyük kentlerin, metropollerin banliyöleri diyebileceğimiz bölgelere hizipleştiklerini görüyoruz. Örneğin İstanbul'da Bağcılar, Esenler bölgesi tarafında ödeklennmeleri olarak ucuz barınma ihtiyaçları, işgücüne kolay erişim bölgeleri olarak görüyoruz. Hem tekstil anlamındaki merdiven altı atölyelerin yoğunluğu, hem de kentsel dönüşüme maruz kalacak daha ucuz

Table 1. *Housing Types of Syrians Living Outside the Camps in Turkey (AFAD 2017, 89)*

Housing Type	Out of Camp		
	Men (%)	Women (%)	Total (%)
Tent	0,90	0,30	0,80
House / Apartment	61,80	65,20	62,40
Temporary Shelter / Plastic / Makeshift	1,20	1,40	1,20
Ruined Building	31,60	30,80	31,50
Public Building	4,3	2,30	4,00
Living in the Street / Open Space	0,20	0,00	0,10
Total (%)	100	100	100
Total Count	1868	351	2.219

As a result of urbanization and high rental prices, many Syrian refugees prefer to live in places that are cheaper and easier to afford, which are typical of houses that are in bad condition and dangerous to live in. Tecer also remarks that a lot of the refugees who live in faulty houses try to fix these issues on their own:

Especially in Bağcılar Istanbul's regions in old settlements, we encountered many families living in basements, shops, places that lack sun and old apartments; we have seen very unsanitary home conditions such as dampness and humidity. The ways of coping of those who stay in half-ruined places and those who pay rent on top of it are through their own makeshift interventions. They take temporary measures, such as pulling a tarp, closing the window with plywood and a board. However, there are no comprehensive shelter projects in our country; shelter is actually the housing issue that creates the most grievance. (Tecer 2021) ³

kiralarn olduğu daha evler, daha altyapısı bulunmayan evlerin bulunduğu bölgeler olduğu karşımıza çıkıyor. Bu, beraberinde çok kötü barınma koşullarını getiriyor. Yani kanalizasyonlarından tutun da evin içerisindeki nem, çatının pencerenin kapının olmadığı evlere şahitlik ettik.”

³ (Turkish Transcription) “Özellikle İstanbul'da Bağcıların çok eski yerleşim yerlerindeki bölgelerde bodrum katlarında, dükkanlarda, güneş görmeye kotlarda, eski apartmanlarda yaşayan birçok aileyle karşılaştık. Rutubet ve nem gibi çok sağlıksız ev koşullarını gördük. Yarı yıkık yerlerde kalanlar ve üstüne bir de kira ödeyen insanların başa çıkma yolları kendilerinin derme çatma müdahaleleriyle oluyor. Bir muşamba çekme, camı kontrplakla bir tahtayla kapatma gibi geçici iş görecekle önlemler alıyorlar. Buna ek olarak bazı projeler de gerçekleştiriliyor sivil toplum kuruluşları tarafından. Ancak çok geniş kapsamlı “shelter projeleri” yok ülkemizde. Shelter aslında en çok mağduriyet yaratan barınma konusu.”

4.3 Finding Housing

There are three main methods to secure shelter that Syrian refugees have been using since their first arrival in Turkey. First, they use familial connections to settle close to them in cities such as Şanlıurfa, Gaziantep, Hatay, Mardin, and Siirt. Additionally, Syrians can find housing services through refugee camps that have been in service for years. The last method is using their own resources to rent or purchase a rental property in any part of Turkey (Taştan et al. 2016, 15). To explain the heavy accumulation of refugees in cities, Turker Saliji talks about Syrian refugees' cycle of arrival to these urban areas, which happens through their social networks:

When some of their acquaintances leave, many more follow them. For a long time, Syrian refugees tried residing in the same house in a very crowded manner (and in certain houses in certain neighborhoods). Unfortunately, their exit from the refugee camps was very uncontrolled. They were able to go wherever they wanted in Turkey, and this created a distorted settlement. When acquaintances, relatives, and people they met (*along the road*) invited more, they began to live in very crowded houses. The following problem still exists: In some cities, the accumulation (*of refugees*) is too high (Istanbul, Hatay, Gaziantep). (*The government*) is starting to get it back under control. In some places, it comes down to not accepting access and directing (*people*) to other cities. Currently, those under temporary protection (if I am registered in Istanbul, I can go to school and use the hospitals located there) can benefit from these services only if they are registered in that city. There are also those who come to Istanbul illegally to benefit. (Saliji 2021) ⁴

It is concluded that the two main reasons Syrian refugees tend to migrate to urban regions are the pre-existing network they have in such areas and easier access to the workforce. Additionally, once they relocate to city centers, Syrian refugees have to deal with issues ranging from urban transformation to life-threatening living conditions.

⁴ (Turkish Transcription) "Bazı tanıdıkları gidince arkasından birçok kişi daha geliyor. Uzun zamandır aynı evin içinde çok kalabalık (belli mahalelerde belli evlerde) barınmaya çalıştılar. Kamplardan çıkmaları maalesef çok kontrolsüz oldu. Türkiye'de istedikleri yere gidebildiler ve bu çarpık bir yerleşim oluşturdu. Orada tanıştıkları kişiler, akrabalar ve arkadaşları da arkalarından çağırınca evlerde çok kalabalık yaşamaya başladılar. Şu problem hala var: bazı şehirlerde yığılmalar çok fazla (İstanbul, Hatay, Gaziantep). Bunu başa dönüp yine kontrol altına almaya başladılar. Bazı yerlerde kabul etmeme ve başka şehirlere göndermeye kadar geliyor. Şu anda geçici koruma kapsamındakiler (eğer ben İstanbul'da kayıtlı isem okula gidebiliyorum, hastaneler kullanabiliyorum) gibi hizmetlerden sadece kimlikleri o şehirde kayıtlıysa yapabiliyor. Kaçak yolla da İstanbul'a gelip yapan olabiliyor."

5. Syrian Refugee Children's Access to Education

5.1 The Turkish Education System

Attending a public or private school for 12 years is mandatory for all boys and girls. The Turkish education system is divided into three stages, or levels. The first level is a four-year primary school (1st, 2nd, 3rd, and 4th grades), the second level is a four-year secondary school (5th, 6th, 7th, and 8th grades), and the third level is a four-year high school (9th, 10th, 11th, and 12th grade) (Fulbright Turkey, n.d.). Secondary schools are not an obligation, but in order to enter Turkish universities, they are a necessity. Primary schools can be public schools, or they can be religiously affiliated. Secondary schools have a variety of options including fine arts high schools, private schools, public high schools, vocational schools, Anatolian high schools, and STEM high schools. Primary and secondary education in public schools is free. Research completed by Eğitim Sen states that the percentage of schools where coeducation is not applied out of schools that accept students by the high school entrance exam (LGS) exceeded 30% in Bolu, Kırşehir, and Siirt and that this rate is 21% in Istanbul. The rate of single-sex, non-co-educational schools in Turkey has reached 15% (T24 2019). Every child who lives in Turkey, regardless of whether they are a citizen, is allowed to benefit from this 12 year-long education system offered in public schools. Thus, any and all under the "temporary protection law" have a right to enroll in one of these public schools (Asylum in Europe 2021).

5.2 Temporary Education Centers

Article 28 of the 2014 Regulation organizes access to education services and stipulates that education will be handled by MEB (Ministry of National Education). Again in 2014, MEB released a "circular on education and teaching services for foreigners". According to this, refugee children would receive education in MEB schools and Temporary Education Centers (TEC), and Provincial-District education commissions would also work there (Bağcı 2020). Temporary Education Centers have been established to provide foreign and refugee students with the opportunity to continue their education. The classes of such centers took place in other school buildings or private spaces. The teachers who are Syrian used Arabic to teach the curriculum which is a slightly modified version of the Syrian curriculum. These centers were connected to the Ministry of National Education. Additionally, Turkish lessons were taught there (Usta et al. 2018, 176). The schooling rates of the students under temporary protection studying in public schools and temporary education centers according to their education levels were 31.63% in kindergarten, 96.90% in primary school, 57.53% in secondary school, and 25.19% in high school. The education level with the highest level of enrollment was reported to be a primary school, and the lowest level of education was high school education (MEB 2018). It is reported that children who are at the high school level have to work because their families are at a low economic level (Emin 2016, 18). In 2017, MEB

called for a retraction of the temporary education centers, and currently, they are not in use anymore (Cumhuriyet 2017).

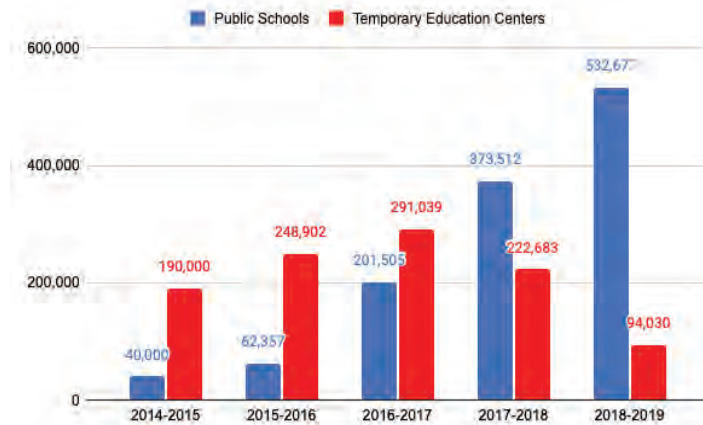


Figure 7. Number of Syrian Students in Public Schools and TECs (2014-2019) (Emin 2019)

5.3 Main Issues That Prevent Syrian Children from Accessing Education

The vast majority of Syrian children in Turkey live outside of refugee camps, in cities and villages, where enrollment rates at schools are much lower. In 2014-2015, only 25% of this population went to school (Human Rights Watch, n.d.).

After UNICEF and the Ministry of National Education prepared an appropriate curriculum, the temporary education centers were terminated, and Syrian students were integrated with Turkish students. For children who lacked Turkish language skills, make-up courses would be offered (Tecer 2021). The effects of urbanization on education such as overcrowded classrooms, underfunded resources, long commute hours, and more posed significant challenges for Syrian children who wish to get an education.

The Turkish education system functions by assigning students (foreigner or resident) to public schools depending on their home address and their proximity to a school. This address-based system can cause problems for immigrant children, whom schools sometimes refuse due to overfilled class sizes; in this case, children are obligated to attend schools that are far away. Living in large cities such as Istanbul can also mean that transportation is highly costly and that additional threats such as kidnapping on the way to school are a reality. Additionally, Tecer explains the role families play in their children's education by highlighting how gossip may affect their decision to send them to school:

Children's families, of course, are the main source of stress and why children do not feel safe on their way to school. Gossip is very widespread amongst the immigrant

community. We even witnessed families taking their children out of school due to violence, verbal abuse, and fear of kidnapping on the way to school. (Tecer 2021) ⁵

Cars driving very fast on the streets and substance addicts who drink alcohol in vacant regions in the middle of a neighborhood can stop the children from going to school and ask for money from them. These reasons also contribute to the road to school being an unsafe place for Syrian children (Büyükhan 2019, 153).

Table 2. *“The child does not go to school because the school path is not safe.” Numerical Distribution of Participants According to Their Answers to the Question (Büyükhan 2019, 151)*

	Number	Percentage (%)
Strongly Disagree	169	66
Disagree	9	3.5
Undecided	0	0
Agree	0	0
Strongly Agree	78	30.5

Reasons such as the early marriage of girls, the need for child labor to supplement the family income, and working as seasonal agricultural workers in rural areas explain why schooling declines as age increases (Tecer 2021). Syrian children have been exploited for labor, which prevents them from being educated. This problem takes place due to the fact that Syrian adults are not given work permits; thus, children are forced to work in illegal workplaces, and in uncontrolled and inhuman conditions (Koman 2015). Saliji also defines children working at jobs in place of their parents as a big obstacle to their education. “Child labor and school dropouts are linked. Children have to work and do not go to school. In general, Syrian refugees work in seasonal jobs as a tribe (Saliji 2021).” ⁶ As a result, they periodically miss class and fall behind in their education. According to the Turkish Statistical Institute (TurkStat), out of 15,247,000 children aged between 6 and 17, 893,000 children were involved in labor with a rate of 5.9% in 2012 (Üstübici and İnce 2020).

While immigrant children typically face obstacles to their education, young girls are at an even bigger risk. Those who wish to get an education

⁵ (Turkish Transcription) “Çocukların aileleri tabii ki en başta stres kaynağı ve çocuklar okul yolunda güvende hissetmiyor kendisini. Dedikodu göçmen toplum arasında çok yayılıyor. Hatta okul yolunda maruz kaldıkları şiddet, sözlü taciz, kaçırılma korkusu gibi olaylardan okuldan çocuklarını alan ailelere şahit olduk.”

⁶ (Turkish Transcription) “Çocuk işçiliği ve okul terkleri bağlantılı. Çocuklar çalışmak zorunda kalıyor ve okula gitmiyor. Genelde Suriyeli mülteciler mevsimlik işlerde çalışıyorlar kabile olarak.”

are often prevented from doing so due to forced marriage as children and subsequent motherhood. In addition, the fact that girls are forced to take on more household chores is seen as one of the biggest reasons for absenteeism from school (Karakaş, n.d.). Based on data released by the United Nations Population Fund (UNFPA), between 2006 and 2017, 13% of Syrian girls under the age of 18 were married (Üstübici and İnce 2020).

Table 3. “The child does not go to school because he/she is engaged/married.” Numerical Distribution of Participants According to Their Answers to the Question (Büyükhhan 2019, 164)

	Number	Percentage (%)
Strongly Disagree	248	96.9
Disagree	0	0
Undecided	0	0
Agree	0	0
Strongly Agree	8	3.1

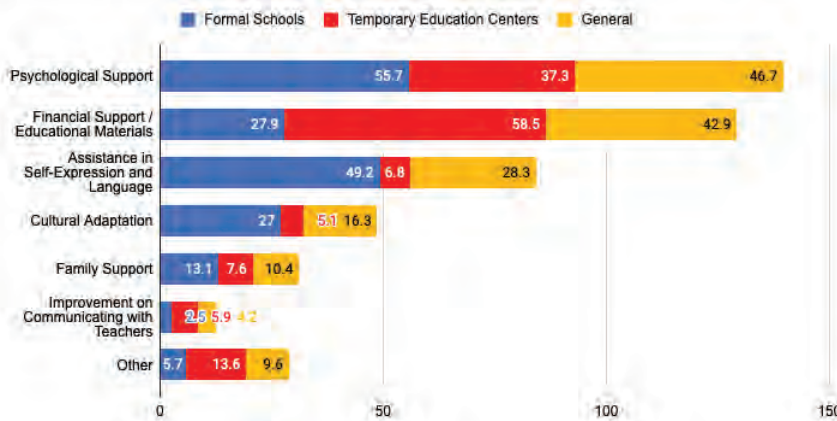


Figure 8. Needs of Syrian Students (%) (2018) (Sonbahar 2019)

6. Services the Government and Organizations Offer to Support Syrian Refugees

Since the beginning of the refugee crisis, the Turkish government and numerous Turkish and international NGOs have been providing comprehensive refugee support services for healthcare, education, housing, safe working conditions, language integration, and many more fields.

6.1 Helping Syrian Refugees With Housing Needs

First or second-degree relatives (even up to third-degree) staying together, large families sharing the same household, 2+ sources of income in a household” (Tecer 2021) are amongst the steps Syrian refugees are taking to cope with the pressures urbanization has put on housing prospects. However, the government or other organizations could attempt to help Syrian refugees find housing through financial or educational means.⁷

Social Cohesion Assistance (SUY/ESSN) provides cash aid to those who live in Turkey under the International Protection / Temporary Protection / Humanitarian Residence Permit and who need help. The program aims to meet the basic needs of asylum seekers and refugees living outside the camps in Turkey, offering food, shelter, and clothing. Aid is given through KIZILAYKART after the assessment of neediness (Kızılay 2017). Ömer Tecer explains that many Syrians use this monetary aid as a way to pay their rent:

(Syrians) are trying to get over the rent issue a little easier by using ESSN assistance. For example, in large families (ex: 2 parents and 4 children), if these children are under the age of 18, the number of this family has enough vulnerability to get ESSN (aid) and 155 Turkish Lira per person. The money coming from here can be sufficient for rental payment. (Tecer 2021)⁸

While this system has been put in place for humanitarian purposes, it may have negative consequences. Some Syrians can develop a negative coping mechanism in order to benefit from the aid to the maximum. The SUY aid can encourage people to have more children, depending on the number of people and the number of children under the age of 18 in the family.

The government also actively tries to accommodate Syrian refugees who are obligated to temporarily visit another city for emergency situations such as for health purposes. Tecer mentions that the economic burden of such visits is lifted by providing basic housing opportunities for the visitors:

⁷ (Turkish Transcription) “Birinci veya ikinci derece akrabaların, hatta üçüncü dereceye kadar gidebilir, birlikte kalmaları, kalabalık ailelerin aynı haneyi paylaşması, bir hane için 1-2-3 gelir kaynağının beslemesi.”

⁸ (Turkish Transcription) “ESSN yardımlarından da yararlanarak kira olayını biraz daha kolay atlatmaya çalışıyorlar. Mesela kalabalık ailelerde (2 ebeveyn 4 çocuk) bu dört çocuk 18 yaşının altındaysa, bu ailenin sayısı ESSN almak için yeterli “vulnerabiliteye” ve kişi başına 155 tl’ye sahip oluyor. Buradan gelecek para kiraya yetebilir.”

Facilities regarding housing opportunities can sometimes be provided as follows: When people visit other cities for compulsory reasons (for health, treatment purposes), they go from one city to another and they have neither acquaintance nor accommodation there (except for hotels). In such cases, if (*the person*) doesn't have the economic opportunity, municipalities, the Ministry of Kızılay and Ministry of Family and Ministry of Social Policy have some shelters; we can think of them as guest houses. If we are to give an example of this from Ankara (In Istanbul, there are those of the municipality so that no one stays out in the cold, especially during winter), opportunities are provided for both foreign and host communities. In the context of illness, the municipalities provide opportunities for the patients' companions. If a foreigner goes somewhere for treatment purposes, they have a social service document that is obtained from the hospital and they can stay in the municipality's shelters thanks to it. The Altındağ, Yenimahalle and Büyükşehir municipalities in Ankara have their own shelter centers and these are called 'compassion houses'. (Tezer 2021) ⁹

In addition to the government, organizations strive to help Syrian refugees when it comes to housing issues as well. For example, 60 unions and NGOs came together in 2012 to form the Şanlıurfa NGO Humanitarian Aid Platform in order to provide food and shelter to homeless or needy Syrians (Türk 2016, 148)

6.2 Helping Syrian Children Have Better Access to Education

The Conditional Education Assistance (CCE) program has been carried out by Turkey since 2003. Over fifteen years, millions of families and children have benefited from this program. In 2017, with the cooperation between the Ministry of Family, Labor and Social Services, the Ministry of National Education, the Turkish Red Crescent, and UNICEF, the National CCE Program was expanded to include school-age refugee children under

⁹ (Turkish Transcription) "Barınma imkanlarıyla ilgili kolaylıklar bazen şöyle sağlanabiliyor: Kişiler zorunlu haller için başka şehirlere gittiklerinde (sağlık, tedavi amacıyla) başka şehirden başka şehire gidiyorlar ve orada ne tanıdığı ne kalacağı yer var (otel hariç). Böyle durumlarda ekonomik imkanı yoksa, belediyeler ve Kızılay ve Aile sosyal politikalar bakanlığının bazı barınma evleri var; misafirhane gibi düşünelim. Ama bunun Ankara'daki örneğini verecek olursak (İstanbul'da da özellikle kış aylarında soğuktan kimse kalmasın diye belediyenin var) hem yabancı hem ev sahibi topluluk olsun imkanlar sağlanıyor. Hastalık bağlamında yine belediyelerin hasta refakatçilerine imkan sağlıyor. Yabancılar da tedavi için geldiyse raporu, hastaneden aldığı bir sosyal hizmet belgesi, oluyor ve bunlarla belediyelerin barınma merkezlerinde kalabiliyorlar. Ankara'da mesela Altındağ, Yenimahalle hem de Büyükşehir belediyesinin kendi barınma merkezleri var ve buna şefkat evleri deniyor."

temporary/international protection in Turkey. To promote education, the beneficiary families whose children continue their education regularly (at least 80% attendance) during the previous school months are provided bi-monthly cash assistance through the Red Crescent Card (UNICEF, n.d.).

Another governmental project, "Project for Supporting the Integration of Syrian Children into the Turkish Education System 2" (PICTES 2) created by the Ministry of National Education (MEB) and the Delegation of the European Union to Turkey, is planned to contribute to Syrian children's access to education and their integration into society (Hürriyet 2018). The project was started on October 3rd, 2016, and still continues in 26 provinces of Turkey (PIKTES, n.d.).

Additionally, for education, social organizations participate in activities to make available better educational opportunities for Syrian refugee children. The multi-purpose community centers established by Uluslararası Mavi Hilal Vakfı (the International Blue Crescent Foundation), SGDD, Malumat, Yuva Derneği (Yuva Association), İnsan Kaynağını Geliştirme Vakıfı [Human Resource Development Foundation (İKGV)], Malumat Toplum Bilgilendirme Merkezi (Malumat Community Information Center), Hayata Destek Derneği (Support to Life Association), and Mavi Kalem Derneği (Mavi Kalem Association) support out-of-school activities for Syrian children (Türk 2016, 151).

IGAM Center for Asylum and Migration Studies is an organization that conducts research on asylum and migration that also strives to ensure that all refugee children and youth of school age, especially young girls, can access education (İGAM, n.d.). With projects like "Education Without Distance", they actively fight to allow Syrian children to have access to English conversation workshops.

MSYD-ASRA (Association of Assistance Solidarity and Support for Refugees and Asylum Seekers), a non-governmental organization based in Ankara, has also completed educational activities for Syrian children such as working with the Sokullu Mehmet Paşa Primary School to complete drama workshops (MSYD-ASRA, n.d.).

6.3 Syrian Self-Help Organizations

There are many Syrian NGOs that also provide public services such as schooling and hospital care as well as shopping opportunities. These are mainly concentrated in the regions with the highest number of Syrians (mainly in certain districts of Istanbul and Gaziantep). These NGOs help Syrians overcome their daily problems, provide legal information, translate from Turkish to Arabic, establish schools and hospitals, and even run newspapers and radio stations (Mackreath and Sağnıç 2017, 59-60). It is important to note that there aren't many publications that talk about or bring attention to such organizations.

7. Recommendations

Many Syrian refugees struggle to keep up with the rapidly rising rental prices, which in turn push them to develop negative coping mechanisms or move to areas that are not sanitary or safe to live in. By identifying the cities where there are high concentrations of refugees, then implementing better rent control policies on such places, municipalities can help provide an economically more stable environment for refugees who already struggle with paying these high rental prices. With tighter rent control regulations, living costs can be more affordable for lower-income Syrian refugees.

While money aid programs such as the Social Cohesion Assistance (SUY/ESSN) aim to give monetary assistance (155 Turkish Lira per person) to refugees under temporary protection, they can also encourage Syrian refugees to have more kids in order to benefit from the system furthermore. Considering that the average rental price in Turkey is 647 TL, while the average rental price in urban cities like Istanbul is 851 TL according to the data of the Turkish Statistical Institute (Milliyet 2017), the government should raise the amount of money per person or supplement it with additional programs in order to overcome this issue.

Syrian children who go to school in urban places have to overcome challenges such as the high cost of transportation, dangers of heavy traffic, and harassment from substance abusers. More traffic workers should be put to work around school districts to control the circulation and ensure the safety of children. In addition, more frequent patrolling of the streets before and after school hours should be enforced by the police. As NGOs and social workers play a crucial role in the integration of Syrian refugees, more cooperation between them and official institutions should happen to identify and find solutions to problems related to housing and education. Programs that aim to educate Syrian children and their parents on the importance of education should be conducted in order to reduce child labor and drop-out rates.

8. Conclusion

Urbanization is a phenomenon that has many negative consequences for housing possibilities and accessing education for Syrian refugees who live in Turkey. In addition to an externally arriving Syrian refugee population, a huge number of Turkish people migrated from villages to urban areas such as big cities as a result of the industrial revolution. This urbanization created many problems; Syrian refugees stayed in places with insufficient infrastructure and accommodated multiple families in one house to afford the high rental prices. These large cities also created problems for children's access to education because of overcrowded classrooms and a high cost of transportation. Additionally, to afford living in cities, children were sent to work instead of going to school.

By interviewing Turkish experts who dedicated their lives to helping refugees, my study adds a new perspective to the struggles of Syrian refugees in Turkey in regards to housing and education as a result of urbanization and

answers the questions “How is urbanization, caused by external and intra-Turkish migration itself, affecting the accessibility of housing prospects and educational opportunities to Syrian immigrants residing in Turkey with an emphasis on large cities like Istanbul?” and “What resources exist to assist this vulnerable community?” Despite new efforts to understand how urbanization influences Syrian refugees’ lives, more research is needed to discover to what extent child marriage and child labor disrupt education, as well as how many self-help organizations there are and what they do to help Syrians in terms of housing and education.

Bibliography

- AFAD. (2017). “Türkiye’deki Suriyelilerin Demografik Görünümü, Yaşam Koşulları ve Gelecek Beklentilerine Yönelik Saha Araştırması [Field Study on Demographic Outlook, Living Conditions, and Future Expectations of Syrians in Turkey].” https://www.afad.gov.tr/kurumlar/afad.gov.tr/25337/xfiles/17a-Turkiye_deki_Suriyelilerin_Demografik_Gorunumu_Yasam_Kosullari_ve_Gelecek_Beklentilerine_Yonelik_Saha_Arastirmasi_2017.pdf.
- Amnesty International. (2016). “Güvenli Olmayan Sığınak: Türkiye’de Sığınmacılar ve Mülteciler Etkili Korunmaya Erişemiyor [Unsafe Shelter: Asylum Seekers and Refugees Can Not Access Effective Protection in Turkey].” <https://www.refworld.org/cgi-bin/texis/vtx/rwmain/opendocpdf.pdf?reldoc=y&docid=575137a84>.
- Aşkın, E. Ö., Yayar R., & Oktay, Z. (2013). “Kırsal Göçün Ekonometrik Analizi: Yeşilyurt İlçesi Örneği [Econometric Analysis of Rural Migration: The Example of Yeşilyurt District].” *C.Ü. İktisadi ve İdari Bilimler Dergisi* 14 (2): 231-252. <https://dergipark.org.tr/tr/download/article-file/48507>.
- Asylum in Europe. (2021). “Country Report: Access to education (Turkey).” Asylum in Europe. <https://asylumineurope.org/reports/country/turkey/content-temporary-protection/employment-and-education/access-education/>.
- Balcı, E. (2020). “Suriyeli Çocukların Eğitimi [Education of Syrian Children].” 21. Yüzyıl Türkiye Enstitüsü. <https://21yyte.org/tr/merkezler/islevsel-arastirma-merkezleri/suriyeli-cocuklarin-egitimi>.
- Bakbak, D. (2018). “A Search into the Container Camps of Syrian Refugees.” *Akademik Bakış* 11, no. 23 (Winter): 249-287. <https://dergipark.org.tr/tr/download/article-file/610381>.
- Bimay, M., & Kaypak, Ş. (2016). “Suriye Savaşı Nedeniyle Yaşanan Göçün Ekonomik ve Sosyo- Kültürel Etkileri: Batman Örneği [Economic and Socio-Cultural Effects of Migration Due to the Syrian War: The Case of Batman].” *Journal of Life Sciences* 6 (1): 84-110. <https://dergipark.org.tr/tr/download/article-file/313312>.

- Büyükhan, M. (2019). "Zorunlu Eğitim Çağındaki Suriyeli Çocukların Eğitime Erişimi Önündeki Engeller: Ankara İli Örneği [Barriers to Access to Education for Syrian Children at the Age of Compulsory Education: The Case of Ankara Province]." (April). http://acikerisim.ybu.edu.tr:8080/xmlui/bitstream/handle/123456789/1193/mithat_buyukhan_tez.pdf?sequence=1&isAllowed=y.
- Çalışkan, A. (2020). "Geçici Koruma Kapsamındaki Suriyelilerin Mekan Tercihlerindeki Rasyonelliğin Analizi [Analysis of the Rationality in the Venue Preferences of Syrians Under Temporary Protection]." *Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 22 (3): 1189-1204. <http://dx.doi.org/10.16953/deusosbil.781822>.
- Cumhuriyet. (2017). "MEB'den Suriyeliler için açılan eğitim merkezlerine kapatma kararı [The decision of the Ministry of National Education to close the training centers opened for Syrians]." *Cumhuriyet*, March 28, 2017. <https://www.cumhuriyet.com.tr/haber/mebden-suriyeliler-icin-acilan-egitim-merkezlerine-kapatma-karari-708584>.
- Emin, M. N. (2016). "Türkiye'deki Suriyeli Çocukların Eğitimi: Temel Eğitim Politikaları [Education of Syrian Children in Turkey: Basic Education Policies]." *SETA*, no. 153 (February). http://file.setav.org/Files/Pdf/20160309195808_turkiyedeki-suriyeli-cocuklarin-egitimi-pdf.pdf.
- Emin, M. N. (2019). "Savaşın Çocuklarını Eğitimle Kazanmak [Winning the Children of War through Education]." *Kriter* 4, no. 38 (September). <https://kriterdergi.com/dosya-turkiyedeki-suriyeli-multeciler/savasin-cocuklarini-egitimle-kazanmak>.
- Erdoğan, M. (2019). "Türkiye'deki Suriyeli Mülteciler [Syrian Refugees in Turkey]." (September). <https://www.kas.de/documents/283907/7339115/Türkiye%27deki+Suriyeliler.pdf/afaf9d37-7035-f37c-4982-c4b18f9b9c8e?version=1.0&t=1571303334464>.
- Erdoğan, M. (2020). "Suriyelilerle Uyum İçinde Yaşamın Çerçevesi [Framework for Living in Harmony with Syrians]." (July). <https://www.unhcr.org/tr/wp-content/uploads/sites/14/2020/09/SB2019-TR-04092020.pdf>.
- Eren, V., & Aydın, A. (2016). "Türkiye'deki Kentleşme Sürecinde Yeni Bir Durum: Suriyeli Kentliler [A New Situation in the Urbanization Process in Turkey: Syrian Urbanites]." *Çukurova Araştırmaları Dergisi* 2, no. 1 (Summer): 31-38. <http://dx.doi.org/10.18560/cukurova.33>.
- Euronews. (2014). "Suriyeli mülteciler Türkiye'ye ağır gelmeye başladı [Syrian refugees have started to be heavy on Turkey]." Euronews. <https://tr.euronews.com/2014/08/02/suriyeli-multeciler-turkiye-ye-agir-gelmeye-basladi>.
- Fulbright Turkey. n.d. "Türk Eğitim Sistemi [Turkish Education System]." Fulbright Türkiye. Accessed July 3, 2021. <https://fulbright.org.tr/turk-egitim-sistemi>.

- Hamsici, M. (2019). "Türkiye'de kaç Suriyeli var, en çok Suriyeli nüfusu hangi şehirde yaşıyor [How many Syrians are there in Turkey, in which city does the largest Syrian population live]?" BBC. <https://www.bbc.com/turkce/haberler-turkiye-49150143>.
- Human Rights Watch. n.d. "Türkiye'deki Suriyeli Mülteci Çocukların Eğitime Erişiminin Önündeki Engeller – Kayıp Nesil Olmalarını Önlemek [Barriers to Syrian Refugee Children's Access to Education in Turkey – Preventing them from Becoming a Lost Generation]." Human Rights Watch. Accessed July 2, 2021. <https://www.hrw.org/tr/report/2015/11/09/282910>.
- Hürriyet. (2018). "Suriyeli çocukların eğitimi için 400 milyon avro [400 million euros for the education of Syrian children]." *Hürriyet*, 12 20, 2018. <https://www.hurriyet.com.tr/egitim/suriyeli-cocuklarin-egitimi-icin-400-milyon-avro-41058615>.
- İçduygu, A., & Aksel, D. B. (2013). "Turkish Migration Policies: A Critical Historical Retrospective." *PERCEPTIONS* 18, no. 3 (Autumn): 167-190. <https://dergipark.org.tr/en/download/article-file/816056>.
- İGAM. n.d. "Neler Yapıyoruz [What We Do]." Accessed July 6, 2021. <https://igamder.org/TR/hakkimizda//neler-yapiyoruz>.
- İnan, C. E. (2016). "Türkiye'de Göç Politikaları: İskân Kanunları Üzerinden Bir İnceleme [Migration Policies in Turkey: A Study on Settlement Laws]." *Göç Araştırmaları Dergisi* 2 (3): 10-33. <https://gam.gov.tr/files/2-Canan-Emek-Inan.pdf>.
- İşık, Ş. (2005). "Urbanisation and Urbanisation Models in Turkey." *Ege Coğrafya Dergisi* 14:57-71. <https://dergipark.org.tr/en/download/article-file/56792>.
- Karakaş, B. n.d. "Türkiye'de 10 Suriyeli çocuktan 4'ü okula gidemiyor [4 out of 10 Syrian children in Turkey cannot go to school]." *DW* (Istanbul). Accessed July 5, 2021. <https://www.dw.com/tr/türkiyede-10-suriyeli-çocuktan-4ü-okula-gidemiyor/a-42353006>.
- Kızılay. (2017). Kızılay Kart. <https://kizilaykart.org/tr/faq0.html>.
- Koman, E. (2015). "Suriyeli mülteci çocuklar ve engellenen eğitim hakkı [Syrian refugee children and the denied right to education]." Heinrich Böll Stiftung Derneği Türkiye Temsilciliği. <https://tr.boell.org/tr/2015/12/29/suriyeli-multeci-cocuklar-ve-engellenen-egitim-hakki>.
- Library of Congress. (2020). "Refugee Law and Policy: Turkey." *Library of Congress Law*, 2020. https://www.loc.gov/law/help/refugee-law/turkey.php#skip_menu.
- Mackreath, H., & Sağnıç, Ş. G. (2017). "Türkiye'de Sivil Toplum ve Suriyeli Mülteciler [Civil Society and Syrian Refugees in Turkey]." (March). <https://hyd.org.tr/attachments/article/216/turkiyede-sivil-toplum-ve-suriyeli-multeciler.pdf>.
- Makovsky, A. (2019). "Turkey's Refugee Dilemma." *American Progress*, (March). <https://www.americanprogress.org/issues/security/reports/2019/03/13/467183/turkeys-refugee-dilemma/>.

- MEB. (2018). "Geçici Koruma Kapsamı Altındaki Öğrencilerin Eğitim Hizmetleri [Education Services for Students Under Temporary Protection]." Hayat Boyu Öğrenme MEB. https://hbogm.meb.gov.tr/meb_iys_dosyalar/2018_10/23093037_22-Ekim_-2018__Yinternet_BYIteni.pdf.
- Mersin University, & Doğan, A. E. (2014). "Suriyeli Göçmenlerin Sorunları Çalıştayı Sonuç Raporu [Final Report of the Workshop on Problems of Syrian Migrants]." (October). <http://www.madde14.org/images/b/b0/MersinUnivSuriyeCalistay.pdf>.
- Milliyet. (2017). "İllere Göre Ortalama Kiralık Daire Fiyatları [Average Rental Flat Prices by Province]." Milliyet. <https://www.milliyet.com.tr/emlak/illere-gore-ortalama-kiralik-daire-fiyatları-69428>.
- MSYD-ASRA. n.d. "Sokullu Mehmet Paşa İlköğretim Okulu ve MSYD-ASRA'dan Mülteci Çocuklara Drama Dersleri [Drama Lessons for Refugee Children from Sokullu Mehmet Paşa Primary School and MSYD-ASRA]." MSYD-ASRA (Association of Assistance Solidarity and Support for Refugees and Asylum Seekers). Accessed July 6, 2021. <https://msyd.org/sokullu-mehmet-pasa-ilkogretim-okulu-ve-msyd-asradan-multeci-cocuklara-drama-dersleri/>.
- Mülteciler Derneği. (2021). "Türkiyedeki Suriyeli Sayısı Mayıs 2021 [Number of Syrians in Turkey May 2021]." Mülteciler Derneği. <https://www.google.com/url?q=https://multeciler.org.tr/turkiyedeki-suriyeli-sayisi/&sa=D&source=editors&ust=1623877912153000&usg=AOvVaw3by7jAJpbaxpp-bj-aDqqk>.
- Orsam & Tesev. (2015). "Effects of the Syrian Refugees on Turkey." no. 195 (January). https://www.tesev.org.tr/wp-content/uploads/report_Effects_Of_The_Syrian_Refugees_On_Turkey.pdf.
- Özer, Mehmet O., & Beyazıt, E. (2020). "Kent Kuramları Bağlamında Türkiye'deki Suriyeli Sığınmacılar [Syrian Refugees in Turkey in the Context of Urban Theories]." *Mustafa Kemal University Journal of Social Sciences Institute* 17 (46): 545-562. <https://dergipark.org.tr/tr/download/article-file/1272683>.
- Özüdoğru, Hülya Y., Kan, A., Uslu, L., & Yaman, E. (2018). "Yerel Halkın Suriyelilere Yönelik Tutum Ölçeği Geliştirme Çalışması [Scale Development Study on Local People's Attitude Towards Syrians]." *Sosyal Politika Çalışmaları Dergisi* 40, no. 2 (October): 115-140. <https://dergipark.org.tr/tr/download/article-file/561586>.
- Piktes. n.d. "Piktes Nedir [What is Pictes?]." PİKTES. Accessed July 6, 2021. <https://piktes.gov.tr>.
- Pişkin, M. (2020). "Dünyanın en büyük mülteci kampı artık Türkiye sınırında: Atme Kampı tehdit oluşturuyor [The world's largest refugee camp is now on the Turkish border: Atme Camp poses a threat]." *Sözcü*. <https://www.sozcu.com.tr/2020/dunya/sondakika-dunyanin-en-buyuk-multeci-kampi-artik-turkiye-sinirinda-atme-kampi-tehdit-olusturuyor-5934552/>.

- Salihi, E. (2016). "Suriyeli Göçmenler Hakkında Yazılan Uluslararası Raporlara Dair Bir Değerlendirme [An Evaluation of International Reports on Syrian Migrants]." *Marmara Journal of Communication* -, no. 25 : 159-164. [10.17829/midr.20162520724](https://doi.org/10.17829/midr.20162520724).
- Salihi, T., interview by author, distant (online), 25 June, 2021.
- Sarmini, I., Topçu, E., & Scharbrodt, O. (2020). "Integrating Syrian refugee children in Turkey: The role of Turkish language skills (A case study in Gaziantep)." *International Journal of Educational Research Open* 1. <https://doi.org/10.1016/j.ijedro.2020.100007>.
- Şentürk, C. (2020). "Türkiye Şehirlerinde Suriyeli Göçmenler [Syrian Immigrants in Turkish Cities]." https://cdn.website-editor.net/09fe2713f5da44ff99ead273b339f17d/files/uploaded/TR-PolicyPaper_8_Syriscche_Migration.pdf.
- Şimşek, D. (2018). "Mülteci Entegrasyonu, Göç Politikaları ve Sosyal Sınıf: Türkiye'deki Suriyeli Mülteciler Örneği [Refugee Integration, Migration Policies and Social Class: The Case of Syrian Refugees in Turkey]." *Sosyal Politika Çalışmaları Dergisi* 18, no. 40/2 (Kasım): 367-392. <https://dergipark.org.tr/tr/download/article-file/590105>.
- Sonbahar, E. (2019). "Suriyeli Çocukların Eğitimi Araştırma Raporu Açıklandı [Research Report on the Education of Syrian Children is Revealed]." Sivil Sayfalar. <https://www.sivilsayfalar.org/2019/03/27/suriyeli-cocuklarin-egitimi-arastirma-raporu-aciklandi/>.
- Stephan, W. G. (2000). "An Integrated Threat Theory of Prejudice." 23-46. https://www.researchgate.net/publication/313054712_An_Integrated_Threat_Theory_of_Prejudice_In_Stuart_Oskamp_ed.
- T24. (2019). "Harem selamlık liseler: "Türkiye'de karma eğitimin uygulanmadığı okul oranı yüzde 15'e ulaştı" [Harem Selamlık high schools: "The rate of schools where coeducation is not implemented in Turkey has reached 15 percent"]." *T24*, August 2, 2019. <https://t24.com.tr/haber/haremlik-selamlık-liseler-türkiye-de-karma-egitimin-uygulanmadigi-okul-orani-yuzde-15-e-ulasti,833179>.
- Taş, H. Y., & Özcan, S. (2018). "Suriyeli Göçmen Sorunlarının, Sosyal Politikalar Bağlamındaki Analizi [Analysis of Syrian Migrant Problems in the Context of Social Policies]." *Emek ve Toplum* 7, no. 17 (January): 36-54. [10.31199/hakisderg.409860](https://doi.org/10.31199/hakisderg.409860).
- Taşkan, C., Kavuncu, A. Ç., Akbulut, A. M., İrdem, İ., Osmanoğlu, E., Özkaya, Ö., & Yıldız, A. K. (2017). "Mass Immigration and Syrians in Turkey." (November). <https://www.pa.edu.tr/Upload/editor/files/1-Mass%20Immigration%20and%20Syrians%20in%20Turkey-DIZGI-TASARIM.pdf>.
- Taşkan, C., Kavuncu, A. Ç., Akbulut, A. M., İrdem, İ., Osmanoğlu, E., Özkaya, Ö., & Yıldız, A. K. (2016). "Uluslararası Kitleli Göçler ve Türkiye'deki Suriyeliler [International Mass Migration and Syrians in Turkey]." no. 5 (November). <https://www.pa.edu.tr/Upload/editor/files/goc-konferansi.pdf>.
- Tecer, Ö. Ç., interview by author, distant (online), (25 June, 2021).

- Terzi, E., & Koçak, Y. (2012). "Türkiye'de Göç Olgusu, Göç Edenlerin Kentlere Olan Etkileri ve Çözüm Önerileri [The Phenomenon of Migration in Turkey, The Effects of Immigrants on Cities, and Solution Suggestions]." *Journal of Economics and Administrative Sciences Faculty* 3 (3). <https://www.kafkas.edu.tr/dosyalar/iibfdergi/file/03/7.pdf>.
- Tokay, M. (2020). "Araştırma: İstanbul'da Suriyeli sığınmacılara yönelik tehdit algıları kemikleşiyor [Research: Threat perceptions towards Syrian refugees are getting stronger in Istanbul]." Euro News. <https://tr.euronews.com/2020/06/12/arast-rma-istanbul-da-suriyeli-s-g-nmac-lara-yonelik-tehdit-alg-lar-kemikleşiyor>.
- TÜİK. (2021). "Adrese Dayalı Nüfus Kayıt Sistemi Sonuçları, 2020 [Address Based Population Registration System Results, 2020]." TÜİK. <https://data.tuik.gov.tr/Bulten/Index?p=Adrese-Dayali-Nufus-Kayit-Sistemi-Sonuclari-2020-37210>.
- Türk, G. D. (2016). "Türkiye'de Suriyeli Mültecilere Yönelik Sivil Toplum Kuruluşlarının Faaliyetlerine İlişkin Bir Değerlendirme [An Evaluation of the Activities of Non-Governmental Organizations for Syrian Refugees in Turkey]." *Marmara Journal of Communication*, no. 25, 145-157. 10.17829/midr. 20162520723.
- UNHCR. (2017). "Türkiye'deki Suriyeli Mülteciler Sıkça Sorulan Sorular [Frequently Asked Questions on Syrian Refugees in Turkey]." (January). <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwih9SJrbDxAhXO4KQKHbedBF8QFjABegQIBBAD&url=https%3A%2F%2Fdata2.unhcr.org%2Fen%2Fdocuments%2Fdownload%2F59167&usg=AOvVawO9BWZ5TP57wAFPL2me3ADr>.
- UNICEF. n.d. "Şartlı Eğitim Yardımı Programı (Ş.E.Y) Şartlı Eğitim Yardımı Programı'nın Suriyeli ve Diğer Mültecilere Yaygınlaştırılması [Dissemination of Conditional Education Assistance Program (Ş.E.Y.) to Syrians and Other Refugees]." UNICEF. Accessed July 6, 2021. <https://www.unicef.org/turkey/sartli-egitim-yardimi-programi-sey>.
- Usta, M. E., Arıkan, İ., Şahin, Y., & Çetin, M. S. (2018). "Suriyeli Eğitimcilerin Gözünden Geçici Eğitim Merkezlerinde Karşılaşılan Sorunların İncelenmesi [Examination of the Problems Encountered in Temporary Education Centers from the Perspective of Syrian Educators]." *Mukaddime*, 173-188. 10.19059/mukaddime. 341921.
- Üstübcü, A., & İnce, H. B. (2020). "Syrian Refugees in Turkey: What Existing Data Implies for Gender and Displacement." Koç KAM. <https://kockam.ku.edu.tr/syrian-refugees-in-turkey-what-existing-data-implies-for-gender-and-displacement-aysen-ustubici-h-berra-ince/>.
- Yıldırım, A. (2004). "Kentleşme ve Kentleşme Sürecinde [Urbanization and Urbanization Process]." <https://dspace.ankara.edu.tr/xmlui/bitstream/handle/20.500.12575/27836/847.pdf?sequence=1>.

- Yılmaz, M. n.d. "Changes of Rural Population in Turkey and its Distribution By Provinces." *Eastern Geographical Review*, 161-188. Accessed June 27, 2021. <https://dergipark.org.tr/tr/download/article-file/27061>.
- Zengin, E. Ç. (2018). "Kent ve Kentleşme Sarmalında Türkiye [Turkey in the Urban and Urbanization Spiral]." *Kastamonu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, (January), 85-103. 10.21180/kuiibf.2018136759.



A Proposal to Facilitate Endogenous, Physiological Adrenal Steroidogenesis in 21-OHD CAH Patients via Recurrent Enzyme Replacement Therapy

Isaac Lowry

Author Background: *Isaac Lowry grew up in the United States and currently attends Granville High School in Granville, Ohio, in the United States. His Pioneer research concentration was in the field of biology and titled “Cell Biology.”*

Abstract

21-OHD CAH is an autosomal recessive disorder that results in impaired function of the 21-hydroxylase enzyme. Enzymatic dysfunction prevents biosynthesis of corticosteroids within the adrenal glands, which, in turn, leads to the accumulation of substrates. These steroid precursors are then rerouted to downstream hormone pathways, causing excessive adrenal androgen production. Modern treatment plans for 21-OHD rely upon oral glucocorticoids and mineralocorticoids, which are intended to address the steroid deficiency and suppress adrenal androgens. However, in order to adequately suppress androgen production, patients often require supraphysiological doses of glucocorticoids. Therefore, striking a balance between managing glucocorticoid excess and achieving optimal androgen suppression is a challenge in modern clinical settings. Patients who sustain chronic glucocorticoid excess and/or chronic hyperandrogenism face a number of health complications, and these complications then necessitate the use of additional, potent medications, such as GnRH analogs, aromatase inhibitors, and growth hormone. The most frequent health problems associated with 21-OHD and its corresponding therapies include low predicted final adult height, obesity, insulin resistance, acute adrenal crisis, cystic acne, elevated blood pressure, increased CIMT measurements, infertility, and reduced quality of life.

Due to the inherent difficulty of managing 21-OHD via conventional therapeutics, this paper will propose the use of recurrent enzyme replacement therapy (ERT) as a treatment solution. ERT involves the intravenous infusion of recombinant enzymes and is currently approved for the treatment of mucopolysaccharidosis, which, like 21-OHD, is characterized by enzyme dysfunction. Recombinant 21-hydroxylase enzymes will be acquired by transfecting the pGFP/DHFR/CYP21A2

plasmid vector into CHO-DG44 cells with subsequent MTX amplification in order to isolate a stable, high-producing cell line. After undergoing two post-transcriptional modifications, the recombinant 21-hydroxylase enzymes will then be delivered to 21-OHD patients on a weekly basis via intravenous infusion. For the purposes of this experiment, frequent blood tests will be collected, and results will be collated into 24-hour profiles for relevant hormones: cortisol, ACTH, 17-hydroxyprogesterone, and androstenedione. Hypothetical results suggest markedly improved hormone profiles for patients receiving ERT compared to patients receiving conventional hormone replacement therapy. This improved hormone regulation would also theoretically confer improved health outcomes, decreased mortality, and improved quality of life for 21-OHD patients.

1. Introduction

Congenital adrenal hyperplasia (CAH) refers to a group of autosomal recessive disorders that result in reduced corticosteroid synthesis by the adrenal glands. This decrease in corticosteroid production stems from the lack of a single necessary enzyme, and over 95% of CAH cases are caused by a 21-hydroxylase deficiency (21-OHD). 21-hydroxylase is responsible for adrenal steroidogenesis as it converts progesterone and 17-hydroxyprogesterone (17-OHP) to deoxycorticosterone and 11-deoxycortisol, respectively (Trapp & Oberfield, 2012). 21-OHD CAH is further categorized as either classic or non-classic. Classic 21-OHD is considered more severe and is typically diagnosed at birth due to the presence of virilized (overdeveloped or ambiguous) genitalia; these children have insufficient cortisol and aldosterone production, as well as excess androgens. Non-classic CAH is less severe, and it often goes undiagnosed until puberty and results in decreased final adult height. Currently, 21-OHD cannot be cured as enzyme dysfunction is caused by a mutation in the CYP21A2 gene. Existing therapies involve hormone replacement therapy (hydrocortisone and fludrocortisone), aromatase inhibitors, gonadotropin releasing hormone analogs, and growth hormone. Patients with 21-OHD, even when receiving treatment, can experience chronic hyperandrogenism as a result of enzyme dysfunction and precursor buildup. 21-OHD is associated with a number of health issues; complications with metabolic, cardiovascular, reproductive, and bone health have all been linked to poor hormone control in CAH patients. Moreover, quality of life for some 21-OHD patients is suboptimal due to difficulty in managing the disorder via modern therapeutics. Current therapies for 21-OHP can cause adverse health effects in patients due to the difficulty of accurately dosing and monitoring medications. (Reisch et al., 2011).

At this point in time, direct enzyme replacement therapy (ERT) has only been used to treat certain lysosomal storage disorders that have poor prognosis with traditional therapies. Recombinant DNA technology is used to produce the enzymes in human fibroblasts, other animal cells, or plant cells, and the enzymes are then delivered intermittently via intravenous infusions. Infusions last 3-4 hours, and they are done every 1 to 2 weeks either in a hospital setting, or, in some

instances, in the patient's home. ERT shows promising results in treating lysosomal storage disorders; urinary GAG, joint health, cardiovascular health, pulmonary function, and quality of life are significantly improved (Concolino et al., 2018). Currently, there is no existing research on implementing ERT for treatment of 21-OHD; instead, research is primarily focused on modified-release glucocorticoid medications and subcutaneous infusion of glucocorticoids. However, these techniques are not yet able to replicate physiological hormone control (Reisch et al., 2011). The aim of this paper is to propose using direct enzyme replacement therapy for 21-OHD patients exhibiting poor hormone control so as to reduce the number of associated health complications and improve quality of life. Increasing the bioavailability of functional 21-hydroxylase enzymes via ERT could theoretically allow patients to resume normal adrenal steroidogenesis, which would resolve chronic hyperandrogenism, hypo- or hypercortisolism, and remove the need for additional medications. Moreover, the frequency of adverse health effects and the risk of mortality would, in all probability, decrease as the patient's hormonal physiology is restored.

2. Basics of Congenital Adrenal Hyperplasia Caused by 21-hydroxylase Deficiency

2.1 What is CAH?

CAH denotes a large family of inherited genetic disorders that cause impaired cortisol production and the overexpression of adrenal androgens. The most common form of the disorder is the 21-hydroxylase deficient (21-OHD) variant, which accounts for over 95% of CAH cases. The classic version of the 21-OHD disorder is considered more severe and can be either simple virilizing or salt-wasting. Classic 21-OHD is associated with having 0-5% enzymatic function, and it is most commonly diagnosed at birth due to the presence of enlarged or ambiguous genitalia in infants. Approximately three out of four individuals with classic 21-OHD present with the salt-wasting form, meaning that — in addition to a decrease in cortisol production — they also have a deficiency in aldosterone production, which impairs their ability to maintain electrolyte homeostasis, putting them at an elevated risk for adrenal crisis. Those with the simple-virilizing form of classic 21-OHD have decreased cortisol production and excess androgens. The non-classic form of 21-OHD is less severe, and patients with this form are characterized as having 10-20% enzymatic function (Speiser et al., 1992). Therefore, patients with the non-classic form often go undiagnosed during infancy, and it is not until early childhood that signs of hyperandrogenism, such as precocious puberty, lead to a diagnosis. Late diagnosis of 21-OHD leads to decreased predicted adult height due to early epiphyseal fusion and growth arrest.

The incidence for the classic form of the disorder is 1:13,000 to 1:15,000 births, whereas the incidence of the non-classic form is much greater (Wilson et al., 2007). Non-classic 21-OHD occurs in roughly 0.1% of newborns in the wider Caucasian population, though it affects a much larger proportion of Ashkenazi Jews (3.7%), Hispanics (1.9%), and Yugoslavs (1.6%). Non-classic 21-OHD is

thought to be the most prevalent autosomal recessive disorder based on global screening results (Speiser et al., 1985).

2.2 Genetic basis of 21-OHD CAH

21-OHD CAH is caused by a mutation in the CYP21A2 gene. The CYP21A2 gene is nearly homologous to its adjacent pseudogene, CYP21P1, and both genes are found on chromosome 6. Each gene contains 10 exons with a high degree of homology, and mutations leading to 21-OHD are the result of unequal crossing over or conversion events between the CYP21A2 and CYP21P1 genes (Marumudi et al., 2013). Almost all mutations are deletions in the CYP21A2 gene or transfers of deleterious nucleotide sequences from the CYP21P1 gene. Of over 100 known mutations in the CYP21A2 gene, 90% of 21-OHD cases are caused by the 10 most common mutations. Disease phenotype — classic salt-wasting, classic simple-virilizing, or non-classic — depends largely on the type and location of the mutation (Speiser et al., 1992). Due to the genetic nature of the disorder, there are currently no cures or permanent solutions for 21-OHD; however, gene therapy has shown strong potential for treating similar, monogenic disorders. Technical and ethical limitations have thus far prevented human trials for gene therapies for 21-OHD; however, due to its high prevalence and difficult management, 21-OHD is a strong candidate for future studies (Trapp & Oberfield, 2012).

2.3 Affected molecular pathways in 21-OHD CAH

2.3.1 Impaired adrenal steroidogenesis

The most immediate effect of impaired 21-hydroxylase function is the inability to synthesize corticosteroids in the adrenal gland. Cortisol and aldosterone are corticosteroids that are both synthesized *de novo* from a common precursor, cholesterol. The process of physiological adrenal steroidogenesis involves several transport proteins and enzymes that allow for the conversion of cholesterol to several precursor hormones, then cortisol and aldosterone.

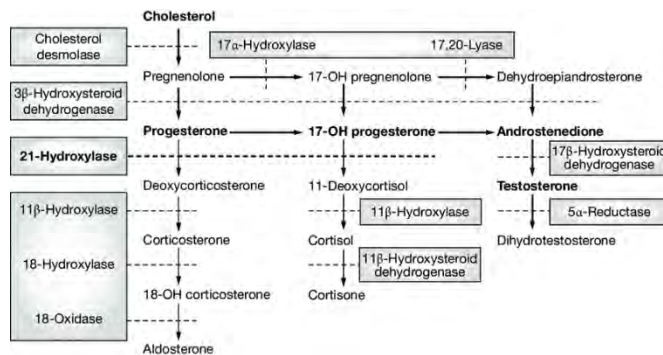


Figure 1. Adrenal steroidogenesis pathways (Nimkarn & New, 2007)

Specifically, the 21-hydroxylase enzyme is essential for the conversion of progesterone to deoxycorticosterone and 17-hydroxyprogesterone (17-OHP) to 11-deoxycortisol. Patients with 21-OHD have severe impairment of this key enzyme, and therefore they cannot produce sufficient cortisol and, in salt-wasting patients, aldosterone. This deficiency can lead to acute adrenal crisis if hormone replacement therapy is not administered (Nimkarn & New, 2007).

2.3.2 Excess adrenal androgen biosynthesis

In addition to leading to a deficiency in corticosteroids, 21-OHD leads to the buildup of precursor hormones that are eventually rerouted to downstream pathways for androgen biosynthesis. 17-OHP is the main substrate of the 21-hydroxylase enzyme, and it is commonly used to diagnose and monitor hormone control in patients with CAH. A buildup of progesterone is also caused by 21-OHD, but it is converted to 17-OHP by 17α -hydroxylase. The excessive amount of intra-adrenal 17-OHP leads to elevated levels of androstenedione, which is then converted to testosterone and later to dihydrotestosterone (DHT). Typically, healthy individuals have levels of 17-OHP below 200 ng/dL; though, patients with the classic form of 21-OHD present with 17-OHP levels exceeding 10,000 ng/dL. Patients with non-classic 21-OHD typically have 17-OHP levels greater than 1,200 ng/dL subsequent to adrenocorticotropic hormone (ACTH) stimulation testing, but levels can often exceed 10,000 ng/dL (Turcu & Auchus, 2015).

2.3.3 Hypothalamic-pituitary-adrenal axis

Reduced 21-hydroxylase function not only induces excessive androgen synthesis via redirected precursor hormones, but it also leads to the dysfunction of the hypothalamic-pituitary-adrenal (HPA) axis as a result of a cortisol insufficiency. This dysfunction is caused by impaired signaling within the negative feedback system between corticotropin-releasing-hormone (CRH), ACTH, and serum cortisol concentrations.

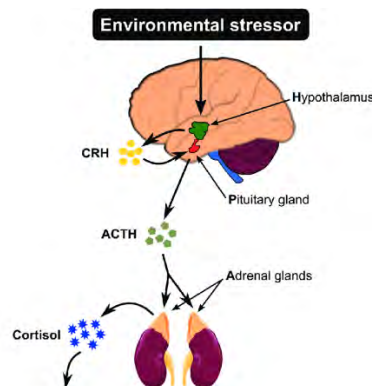


Figure 2. Hypothalamic-pituitary-adrenal (HPA) axis (Lanoix & Plusquellec, 2013)

This absence of proper signaling, in turn, causes an excessive stress response by the hypothalamus and pituitary gland in patients with untreated adrenal insufficiency. CRH and ACTH are produced in surplus due to the lack of appropriate cortisol synthesis, which acts as a termination signal for stress responses. Overstimulation of ACTH further contributes to hyperandrogenism in 21-OHD patients since ACTH primarily acts upon the adrenal cortex and leads to the overproduction of adrenal androgens (Chrousos et al., 2009). The impact of cortisol on the HPA axis explains why continuous monitoring and redosing for CAH patients is paramount (and difficult), as having cortisol levels above or below normal leads to inappropriate degrees of ACTH stimulation.

2.4 Clinical implications of 21-OHD

Patients with 21-OHD often suffer from several health complications; they are either caused by the disorder itself or iatrogenic (caused by medical treatment). Hyperandrogenism mainly contributes to health problems in CAH patients at a young age, or even prior to birth. Neonatal hyperandrogenism causes virilization in patients with the classic form of the disorder, which leads to enlarged or ambiguous genitalia in male and female patients, respectively. Furthermore, patients with non-classic 21-OHD oftentimes go undiagnosed during infancy, and those patients will likely face precocious puberty due to chronic hyperandrogenism during their childhood; this early puberty induces accelerated growth velocity at a young age coupled with advanced bone aging. Advanced bone age is defined as when the degree of maturation of a patient's bones exceeds their chronological age. Advanced bone aging causes early epiphyseal (growth plate) fusion, which often results in patients having a lower predicted height compared to their mid-parental target height (Kamoun, 2013). Hyperandrogenism, excess/reduced cortisol, and excess/reduced aldosterone can also result in decreased bone, metabolic, and cardiovascular health; less common effects include infertility, impaired psychological well-being, and reduced health-related quality of life (HRQoL). Nearly all symptoms in 21-OHD patients are the result of poor hormone control; even during the course of treatment for the disorder, patients can exhibit poor hormone control as modern therapies do not address the underlying cause of the disorder. Thus, dosing and monitoring treatment in 21-OHD patients are oftentimes suboptimal, as all individual patients have different requirements and respond differently to existing medications (Reisch et al., 2011).

3. Existing therapies for patients with 21-OHD CAH

3.1 Glucocorticoid replacement therapy

The hallmark treatment for 21-OHD is the supplementation of glucocorticoids via oral hormone replacement therapy. The most common drug used in adrenal insufficiency patients is hydrocortisone (Cortef); though, prednisone and dexamethasone are also used in some adult patients. Physiological levels of cortisol production are thought to be 5-6 mg/m²/d. However, in 21-OHD patients, typical glucocorticoid replacement therapy involves 15-25 mg/m²/d,

divided between two or three daily doses, for it is not meant solely as a minimal replacement regimen; glucocorticoids also serve to suppress adrenal androgen production. Therefore, adequate cortisol replacement is not synonymous with complete hormone control, and, in numerous cases, achieving full suppression of androgens and 17-OHP is only possible via supraphysiological doses of glucocorticoids (Kamoun, 2013). For this reason, glucocorticoid therapy — although proven to help 21-OHD patients to lead normal lives — is not entirely ideal, as overtreatment with corticosteroids can cause growth suppression and other metabolic syndrome effects, whereas undertreatment leads to hyperandrogenism. Therefore, continuous monitoring of 17-OHP, androstenedione, height, weight, and blood pressure is essential for maintaining an appropriate dosage of glucocorticoids. Hydrocortisone (Cortef) is the preferred oral glucocorticoid drug for adolescents with 21-OHD due to its short, eight-hour half-life, which decreases the risk of growth suppression compared to longer-acting medications. However, many adult patients with 21-OHD elect to switch to a single daily dose of prednisone — a glucocorticoid with a longer biological half-life — subsequent to growth cessation (Ahmed et al., 2019).

3.2 Mineralocorticoid replacement therapy

Oral fludrocortisone (Florinef) is used to replace mineralocorticoids in salt-wasting 21-OHD patients that exhibit a deficiency in aldosterone. Patients with the salt-wasting form of classic 21-OHD have complete dysfunction of aldosterone biosynthesis, which leads to a deficiency in salt (hyponatremia), elevated potassium (hyperkalemia), low blood pressure, and dehydration — all characteristics of adrenal crisis. Hence, patients with salt-wasting 21-OHD receive fludrocortisone supplementation during infancy in conjunction with typical glucocorticoid replacement; it is also recommended that salt-wasting infants receive 5-8 mmol/kg/day of sodium chloride supplementation due to high mineralocorticoid resistance in newborns (Gomes et al., 2013). Typical dosage for fludrocortisone replacement ranges from 200 µg per day in the first six months of life to 125 µg per day after two years of age. Careful monitoring of blood pressure, sodium, potassium, and plasma renin levels is essential for making dose adjustments for fludrocortisone as overtreatment with fludrocortisone can cause hypertension, obesity, and insulin resistance. Once the patient reaches adolescence and adulthood, they generally require a maximum dose of 100 µg per day (Padidela & Hindmarsh, 2010).

3.3 Off-label drug usage in 21-OHD patients

3.3.1 Gonadotropin Releasing Hormone Analog (GnRHa)

As mentioned in previous sections, accelerated bone aging is a concern in patients with non-classic 21-OHD that go undiagnosed during infancy, since accelerated bone aging causes early epiphyseal fusion. Patients that go undiagnosed until later stages of life are normally identified after signs of hyperandrogenism become evident, such as central precocious puberty, increased growth velocity, and tall stature during adolescence. Many patients with non-classic 21-OHD do

not present symptoms of cortisol deficiency; instead, they only show signs of excess androgens (Trapp & Oberfield, 2012). Moreover, patients with classic 21-OHD — both the simple virilizing and salt-wasting form — can similarly experience central precocious puberty and accelerated bone aging due to chronic, mild-to-moderate hyperandrogenism, which is typical even when hormone replacement therapy has been administered since infancy. A discrepancy between chronological age and bone age causes poor height outcomes in 21-OHD patients. One method of counteracting accelerated bone age is through gonadotropin releasing hormone analogs (GnRHa), which suppress estrogen and testosterone by inhibiting the synthesis of luteinizing hormone (LH) and follicular stimulating hormone (FSH). A study tracking pediatric 21-OHD patients found that GnRHa-treatment was effective in decreasing the disparity between bone age and chronological age. However, many patients who undergo GnRHa treatment experience a decrease in growth velocity, so their predicted adult height remains suboptimal; therefore, it is not uncommon for 21-OHD patients to receive GnRHa and growth hormone simultaneously in order to preserve pre-treatment height velocities (Güven, 2014). Frequently appearing side effects of GnRHa therapy include fluid retention, weight gain, gynecomastia, fatigue, and metabolic abnormalities such as the worsening of osteoporosis (decreased bone mass density) and diabetes (LiverTox, 2018).

3.3.2 Aromatase inhibitors

Another novel method of reducing the rate of epiphyseal fusion — which causes growth arrest and short stature in 21-OHD patients — is through the suppression of estradiol. Estradiol is a primary regulator of epiphyseal fusion due to its role in promoting senescence of progenitor cells in chondrocytes (cells which are found at the ends of epiphyseal plates). Aromatase is an enzyme responsible for converting testosterone to estradiol. Patients with 21-OHD can have elevated levels of testosterone due to excessive androgen biosynthesis, which in turn leads to elevated levels of estradiol. Thus, aromatase inhibitors (Anastrozole) are used in 21-OHD patients that have severe bone age acceleration, such that the rate of bone maturation is slowed while maintaining normal height velocity, thus increasing predicted adult height. Aromatase inhibitors have been shown to independently increase predicted adult height when used alongside conventional glucocorticoid and mineralocorticoid replacement therapy. However, individual case studies of CAH patients have shown that aromatase inhibitors, when combined with growth hormone therapy and/or GnRHa, have a greater impact on final adult height outcomes. Furthermore, a study of children with CAH receiving aromatase inhibitors alongside GnRHa (without growth hormone therapy) showed a significant increase in predicted height versus patients receiving only GnRHa. Aromatase inhibitors have been proven to achieve remarkably improved height outcomes in patients with bone age advancement exceeding seven years over chronological age; however, 21-OHD patients who use aromatase inhibitors can frequently experience osteoporosis, joint pain, gynecomastia in men, and menopausal symptoms in women, such as hot flashes and night sweats due to poor hormone balance during treatment (Goedegebuure & Hokken-Koelega, 2019).

3.3.3 - Growth hormone (GH) therapy

Patients with 21-OHD CAH — even with excellent hormone control throughout childhood and puberty — commonly have low final adult heights. Glucocorticoid therapy can suppress endogenous growth hormone (GH) production and limit intestinal calcium absorption, which contributes to poor height outcomes in 21-OHD patients. Therefore, GH therapy is often used in conjunction with the aforementioned gonadotropin releasing hormone analogs (GnRHa) and aromatase inhibitors in patients with accelerated bone aging. GH has proven effective in improving height outcomes without the concurrent use of GnRHa or aromatase inhibitors. A study of 20 CAH patients found that both GH therapy and GH-GnRHa combination therapy significantly increased growth velocity and final predicted adult height compared to comparable CAH controls who did not receive GH therapy (Quintos et al., 2001).

4. Health Outcomes for Patients Receiving Conventional Therapies

4.1 Metabolic health

4.1.1 Obesity, increased BMI, and body fat percentage

A major concern for patients receiving treatment for 21-OHD is the risk of increased BMI, obesity, and elevated body fat percentages. Glucocorticoid excess is thought to be the primary cause for adverse changes to body composition via iatrogenic Cushing's syndrome. A study of 89 children with classic 21-OHD found that there was a significantly greater frequency of obesity (BMI standard deviation score (SDS) > 2.0) in CAH patients receiving glucocorticoid therapy compared to normal, age-related populations. There was no significant difference observed between patients receiving different oral glucocorticoids medication (hydrocortisone, prednisone, dexamethasone), nor was there a significant difference between different forms of the disorder (salt-wasting versus simple-virilizing) (Volkl, 2006). However, several other studies indicate that adults and children with CAH have greater body fat and that the degree of adiposity is dependent on the actual glucocorticoid dosage. Moreover, a study of 37 children with CAH found that individuals with the classic form of 21-OHD had greater body fat compared to comparable, age-related controls, and, in this experiment, increased adiposity was attributed to prolonged exposure to suprphysiological doses of glucocorticoids (Williams et al., 2010).

4.1.2 Insulin resistance

21-OHD patients can also experience low insulin sensitivity as a result of undertreatment as well as overtreatment of glucocorticoids. A study of 18 pediatric CAH patients associated insulin resistance with elevated concentrations of serum leptin (a hormone that is secreted by white adipose tissue and regulates energy homeostasis). The serum leptin concentrations of the CAH patients —

even when adjusted for BMI differences — were significantly higher compared to comparable control subjects. The CAH patients also had significantly greater serum insulin levels when fasting, and, using the homeostasis model assessment (HOMA-IR) method, it was determined that CAH patients also showed significantly greater insulin resistance compared to control subjects. Increased serum leptin and insulin, as well as decreased insulin sensitivity, were attributed to long-term glucocorticoid therapy, though varying doses of hydrocortisone did not correlate with leptin or insulin concentrations between CAH patients (Charmandari et al., 2002). Additionally, a separate study of 37 children found that non-classic CAH patients had higher serum insulin concentrations after stimulation, and this was attributed to prolonged postnatal hyperandrogenism. Thus, undertreatment with glucocorticoids is thought to induce negative metabolic effects (i.e., impaired insulin sensitivity), as adequate androgen suppression is not achieved (Williams et al., 2010). The prevalence of insulin resistance in CAH patients once again elucidates the difficulty of treating the disorder with modern hormone replacement therapy as both overtreatment and undertreatment can induce metabolic syndrome effects. Attempting to balance androgen suppression without excessive glucocorticoid is a major problem for today's physicians as negative health effects are largely unavoidable over the course of conventional treatment plans.

4.1.3 Acute adrenal crisis

Perhaps the most concerning health complication associated with 21-OHD is the increased risk of acute adrenal crisis during instances of dangerously low cortisol and aldosterone. Infants born with the salt-wasting form of classic 21-OHD generally have the greatest risk of adrenal crisis, which is oftentimes lethal if glucocorticoid and mineralocorticoid treatment are not initiated. Furthermore, a retrospective study of 588 adults with 21-OHD found that the mean age of death was 41.2 ± 26.9 years in CAH patients and 47.7 ± 27.7 years in controls ($P < .001$). 42% of premature deaths were caused by acute adrenal crisis. Patients with more severe forms of the disorder (i.e., salt-wasting individuals) had the highest risk of mortality. This study establishes 21-OHD as a potentially lethal condition for adults due to adrenal crisis. Increased risk of adrenal crisis is largely attributable to the fact that conventional hormone replacement therapy does not adequately protect patients against the threat of acute adrenal crisis (Oster et al., 2016; Falhammar et al., 2014).

4.2 Cardiovascular risk factors

4.2.1 Elevated blood pressure

Due to the increased risk for above-average BMI in 21-OHD patients, elevated blood pressure is a concern, and it must be closely monitored to reduce cardiovascular risks. A study of 38 children with classic, salt-wasting 21-OHD showed that mean systolic and diastolic blood pressure were significantly higher in 21-OHD patients compared to a reference population. 58% of patients had systolic hypertension, and 24% had diastolic hypertension. Moreover, 84% of 21-

OHD patients did not experience a physiological, nocturnal dip in systolic blood pressure. (Roche et al., 2003). Another study of 11 CAH patients receiving supraphysiological doses of glucocorticoids noticed higher mean systolic and diastolic blood pressure while patients were awake. 63% of patients did not experience a nocturnal dip in systolic blood pressure, and 54% of patients had nocturnal hypertension (de Silva et al., 2004). The absence of a nocturnal dip in systolic blood pressure has been described in several other studies of CAH patients. Notably, a cross-sectional meta-analysis of 20,000 patients was conducted via public databases analysis, and it found that this “non-dipper” phenomenon resulted in a significantly increased risk for adverse cardiovascular events (Gorostidi et al., 2007). Some studies have shown normal or decreased blood pressure in 21-OHD patients, so it is thought that elevated blood pressure is primarily caused by increased BMI rather than the disorder itself. Though, it remains unclear whether glucocorticoid and mineralocorticoid therapies could induce iatrogenic hypertension. Thus, it is important that salt-wasting 21-OHD patients are closely monitored for plasma renin measurements while receiving mineralocorticoid therapy (Reisch et al., 2011).

4.2.2 Increased carotid intima-media thickness measurements

Several past studies have indicated that 21-OHD patients may have an elevated risk of stroke and myocardial infarction due to increased carotid intima-media thickness (CIMT) measurements. CIMT is used as a surrogate marker for atherosclerosis, which is a condition characterized by the accumulation of fatty material plaques on the inner walls of arteries. A meta-analysis of eight studies determined that CIMT was a strong predictor of future vascular events; an absolute difference of 0.1 mm in CIMT measurement increased the risk of myocardial infarction by 10-15% and the risk of stroke by 13-18% (Lorenz et al., 2007). Furthermore, several studies have linked 21-OHD to having significantly increased CIMT. One study found that 19 adult 21-OHD patients had increased average intima-media thickness at all four sites that were tested: right and left carotid bulbs and right and left common carotids. Data in this study also suggested that greater CIMT (and other intima-media thickness sites) was independent of BMI, high blood pressure, or altered lipid profiles (Sartorato et al., 2007). Another study of pediatric 21-OHD found that CIMT correlated with 17-hydroxyprogesterone ($r = 0.48$, $P < 0.05$) and androstenedione ($r = 0.46$, $P < 0.05$); though, total CIMT in these patients did not significantly differ from controls (Kim et al., 2016). This may suggest that increased CIMT measurements in 21-OHD patients are closely related to the degree of hormone control in an individual and that increased CIMT is less common in younger individuals.

4.3 Bone Health

As a consequence of impaired metabolic regulation, supraphysiological glucocorticoid doses, and iatrogenic hypoandrogenism, 21-OHD patients can experience several adverse effects on their bone health. Glucocorticoid treatment is the leading cause of therapy-induced osteoporosis by affecting both osteoblasts and osteoclasts, which are responsible for bone formation and reabsorption, respect-

ively. Chronic glucocorticoid excess leads to increased activity, maturation, and survival of osteoclasts via increased receptor activity of RANKL, a transmembrane glycoprotein expressed on the surface of stromal cells in the bone. Moreover, RANKL activity suppresses apoptosis in mature osteoclasts and enhances apoptosis in osteoblasts, further contributing to a net-negative effect on bone mass density (BMD). These bone-related effects of supraphysiological glucocorticoid dosages were consistent in a cross-sectional study of 30 classic 21-OHD patients and 138 healthy controls; the results of the study showed that, in addition to being shorter than controls, 21-OHD patients had significantly lower height-adjusted, whole-body BMD. The patient's BMD did not significantly correlate to the actual glucocorticoid dose, nor did it correlate with 17-OHP, renin, androstenedione, or testosterone. The same study also found that 21-OHD patients had increased serum concentrations of bone-specific alkaline phosphatase (BALP), a marker of bone formation, and carboxy-terminal telopeptide of type I collagen (CTX), a marker of bone absorption. This suggests that patients likely have a greater bone turnover rate — a condition that markedly increases the risk for osteoporosis. Increased bone turnover rate due to supraphysiological glucocorticoid treatment and impaired regulation of bone metabolism hormones seems to be a reasonable explanation for decreased BMD in 21-OHD patients. In fact, studies of Addison's disease patients — who have a similar deficiency in cortisol, without the characteristic hyperandrogenism of 21-OHD — have shown no significant difference in BMD. This is likely due to the fact that 21-OHD patients receive excess glucocorticoids via hormone replacement therapy to suppress adrenal androgens, whereas Addison's disease patients receive a minimal replacement dose of glucocorticoids. Another theory for decreased BMD in 21-OHD patients is that early epiphyseal fusion could reduce bone accretion during skeletal maturation in both classic and non-classic 21-OHD patients (Sciannamblo et al., 2006; Rehman & Lane, 2003).

4.4 Sex-related health effects

4.4.1 Female-related health complications

Chronic hyperandrogenism in women with late-onset (non-classic) 21-OHD typically results in symptoms similar to those seen in individuals with polycystic ovary syndrome (PCOS). Both disorders are characterized by the presentation of hirsutism (excessive hair growth), menstrual irregularities, male-pattern baldness, and severe cystic acne. Based on current evidence, it is thought that PCOS and non-classic 21-OHD are overdiagnosed and underdiagnosed, respectively (Savlov et al., 2008). Furthermore, in addition to having a similar phenotypic presentation as PCOS, it is also thought that dysregulation of the HPA axis in 21-OHD patients can lead to the development of ovarian cysts, which may cause infertility, insulin resistance, and obesity (New, 1993). A meta-analysis of 97 articles found that 30-40% of women with non-classic 21-OHD were observed to have polycystic ovaries, with some individual studies indicating higher percentages, up to 82%. Non-classic 21-OHD females were also found to have greatly increased risk for metabolic syndromes, obesity, insulin resistance, and mood disturbances as a result of altered gonadal hormone secretion (Papadakis et al., 2019).

4.4.2 Male-related health complications

Chronic hyperandrogenism in males with non-classic 21-OHD typically leads to severe cases of cystic acne that are resistant to topical and systemic medications; moreover, cystic acne is often the only symptom of androgen excess in males with non-classic 21-OHD (Savlov et al., 2008; Placzek et al., 1999). Additionally, testicular adrenal rest tumors (TART) are a major concern in males with 21-OHD as they have been reported in up to 94% of 21-OHD cases. TARTs are thought to develop from aberrant adrenal cells descending during embryogenesis with testicular cells, and the incidence of TARTs does not typically correlate with overall hormone control. Tumors can lead to drastically reduced fertility; though, treatment with dexamethasone has shown promising results for reducing tumor size and the risk of infertility (Reisch et al., 2011). However, additional treatment with dexamethasone would further contribute to glucocorticoid excess, inevitably causing even more health effects in those patients.

4.5 Quality of life effects

21-OHD patients treated with the current standard of care often experience reduced quality of life as a result of chronic glucocorticoid excess, chronic hyperandrogenism, and all related health complications. A questionnaire-based study of 45 pediatric patients found that fatigue scores did not vary between children with 21-OHD and controls, except in the sleep/rest subscale. It is thought that scores did not significantly differ due to the desensitized perception of fatigue in 21-OHD patients (i.e., they may have reduced energy levels, but they do not consider it abnormal since they have had the condition since birth). The significant difference in sleep/rest fatigue is very likely due to the altered circadian rhythm during glucocorticoid replacement therapy, as 21-OHD patients often have higher nocturnal cortisol concentrations compared to physiological levels in healthy controls (Halper et al., 2017). Several studies of 21-OHD patients have found that quality of life is correlated with adiposity, insulin resistance, and body fat distribution. Furthermore, quality of life did not correlate with disease phenotype (non-classic versus classic), which suggests that reduced quality of life is more a consequence of glucocorticoid excess and poor hormone control during conventional treatment. Overall, there is limited scientific literature concerning quality of life outcomes in 21-OHD patients; although, existing studies suggest that most instances of reduced quality of life are due to the limited hormone regulation that is possible with current treatments (Hans et al., 2013; Reisch et al., 2011; Hans et al., 1998).

5. History and Efficacy of Enzyme Replacement Therapy

5.1 Enzyme replacement therapy in clinical practice for lysosomal storage disorders

Over the past two decades enzyme replacement therapy (ERT) has been employed as a treatment for mucopolysaccharidosis (MPS), which refers to a group of

inherited lysosomal storage disorders where an enzyme deficiency causes the buildup of glycosaminoglycans (GAGs) within lysosomes. ERT requires the production of recombinant enzymes within human fibroblasts, plant cells, or animal cells (Chinese hamster ovary (CHO) cells). Recombinant enzymes are administered to patients via periodic intravenous (IV) doses; the enzymes are captured by cells due to the presence of Mannose-6-phosphate (M6P) on recombinant enzymes, which labels them for endocytosis and transport to the lysosome. From there, recombinant enzymes catabolize the accumulated GAGs. ERT infusions occur every 1-2 weeks, and they take 1-4 hours depending on the disease-specific enzyme and patient's individual dose. Most infusions occur in a hospital setting due to the risk of severe anaphylaxis, but at-home ERT is available for certain disorders and is determined based on a risk-benefit analysis by the physician and patient. Currently, ERT is approved for the treatment of MPS I (Hurler, Hurler-Scheie, Scheie syndrome), MPS VI (Maroteaux-Lamy syndrome), MPS II (Hunter syndrome), MPS IVA (Morquio A syndrome), and MPS VII (Sly syndrome) (Concolino et al., 2018).

5.2 Efficacy of ERT

ERT has greatly improved disease prognosis for patients suffering from MPS disorders, and although it does not cure the disorder, it can lead to significant somatic benefits. In the past, supportive, symptom-based treatment has been the standard for multisystem, progressive disorders such as MPS; however, ERT offers a way to treat the underlying enzyme deficiency and has been shown to decrease the need for additional treatments as well as slowing the progression of the disorder. MPS disorders have historically been diagnosed and monitored through urinary GAG concentrations and organomegaly (organ enlargement) measurements. ERT has been shown to dramatically decrease urinary GAG secretion in the first 3-6 months with a subsequent, slow decline during the following years, suggesting a high efficacy for MPS patients. Furthermore, organomegaly (liver and spleen volumes) decreased rapidly during the first few months of ERT. These impressive results have been proven in numerous studies, and many patients have been able to maintain normal urinary GAG and organ size measurements as a result of ERT. Musculoskeletal benefits have also been reported for MPS patients who receive ERT, such as improved range of motion in the shoulders and decreased wrist laxity. Heart inflammation — a common appearance in MPS patients — has been reduced in patients who receive ERT; specifically, myocardial hypertrophy has been proven to be ameliorated in many cases of ERT treatment as a result of improved geometry and contraction of the heart muscle. Health-related quality of life (HRQoL) and activities of daily living are also improved in MPS patients who receive ERT as long as they do not yet have significant cognitive decline prior to beginning treatment (Concolino et al., 2018; Valayannopoulos & Wijburg, 2011.)

5.3 Limitations of ERT

Despite the frequent success of ERT in modern treatment of MPS disorders, there are some limitations and concerns for long-term treatment with this method. The

foremost issue with ERT is the potential for adverse immune responses to the infusion of recombinant enzymes, such as anaphylaxis. However, a 26-week study of 45 patients with MPS I found that recombinant human alpha-l-iduronidase (laronidase) infusions were well tolerated in all but one patient, who suffered from an anaphylactic reaction. 53% of patients experienced mild infusion associated reactions (IARs); however, IAR's were easily managed and markedly subsided subsequent to the first six months of ERT administration (Clarke et al., 2009). IARs are not uncommon and can appear in as many as 90% of MPS IVA patients receiving ERT; they include skin rashes, facial swelling, constriction of the airways, and nasal swelling, as well as very rare instances of moderate or severe anaphylaxis. Desensitization to recombinant enzymes has been indicated in patients with MPS I, MPS II, and MPS VI, thus decreasing the prevalence and severity of IARs over the course of prolonged ERT. Furthermore, immunogenicity — specifically, the formation of anti-drug antibodies (ADAs) — is another concern for the long-term efficacy of ERT. ADAs appear in 50-90% of patients receiving ERT depending on the type of MPS disorder. It is thought that ADAs could be responsible for impaired biological efficacy of the recombinant enzymes via altered enzyme targeting, increased enzyme turnover, and inhibition of the catalytic site. However, it has yet to be determined whether the formation of ADAs decreases the efficacy of ERT over long treatment periods, and thus far studies have failed to demonstrate decreased safety for patients who develop ADAs in response to treatment (Concolino et al., 2018).

A secondary, though extremely important, consideration for ERT is the high cost of the treatment. A retrospective study of 68 patients who received ERT for MPS I determined the average annual cost of treatment with recombinant laronidase enzyme in 2011: \$354,488 for adults, \$191,757 for children, and \$38,257 for infants with the disease. The high treatment cost is mostly attributable to the cost of the recombinant enzymes themselves, as pharmaceutical companies must recoup their immense research expenditures when developing novel therapies for rare disorders. Furthermore, the cost of developing appropriate cell lines, engineering genetic vectors, and purifying recombinant enzymes is very costly (Wyatt et al., 2012). However, the cost of producing therapeutic enzymes could potentially be much lower for other disorders such as 21-OHD due to the relative prevalence being much higher; if more patients were to utilize the medication, then costs would decrease as increased production would lower average expense.

6. Methodology of Proposed Study

6.1 High-level production of recombinant 21-hydroxylase

6.1.1 Overview of recombinant enzyme production

The first step in producing a clinically viable amount of recombinant 21-hydroxylase enzymes in suspension is the preparation of a cell line with adequate levels of protein production. For this experiment, animal cells will be used rather than prokaryotic cells due to the fact that end products do not require refolding, and posttranslational modifications can be more easily added to the final proteins

(Jenkins & Curling, 1994). Specifically, Chinese hamster ovary (CHO) cells will be used since their chromosomes have several transcriptionally active sites, a trait which favors recombinant protein proliferation. CHO cell products also possess greater biological viability and efficacy as therapeutic proteins. However, production of specific proteins typically occurs at a suboptimal rate in CHO cells, thus requiring clones to undergo gene amplification. Identifying an acceptable cell line will require the preparation of CHO-DG44 cells, where deletion of the dihydrofolate reductase (*DHFR*) gene has occurred, followed by the co-transfection of a *CYP21A2* gene alongside a *DHFR* minigene and green fluorescent protein (GFP) sequence, using the same vector. CHO cells which have successfully integrated the expression vector will survive in cell culture media that are deficient of hypoxanthine and thymidine (HT), since the *DHFR* enzyme is necessary for the endogenous biosynthesis of purines. All surviving CHO cells will then be isolated and considered stable clones. In order to identify clones that have undergone gene amplification of the GFP/*DHFR*/*CYP21A2* gene, CHO cells will be exposed to gradually increasing concentrations of methotrexate (MTX); cells that survive in the presence of MTX are typically found to have undergone gene amplification at the integration site (Nunberg et al., 1978). Further selection of high-producing clones will occur subsequent to the first round of MTX exposure, where individual CHO cells will be separated and exposed to higher concentrations of MTX, and the most transcriptionally active clones will be isolated through the use of a fluorescence-activated cell sorters (FACS) and Southern blot analysis. After a high-producing clone has been identified, it will be transferred to a spinner flask with an appropriate culture media; levels of recombinant 21-hydroxylase enzymes will then be measured for the following seven days with an evaluation of productivity. Recombinant 21-hydroxylase enzymes will also require two post-translational modifications — a mannose-type oligosaccharide and nascent polypeptide N-terminal signal sequence strand — in order to mediate cell membrane endocytosis and to properly label them for endoplasmic reticulum (ER) uptake (Cacciatore et al., 2010; Hong et al., 2018; Kito et al., 2002).

6.1.2 Engineering pGFP/*DHFR*/*CYP21A2* plasmid vector

Effective vectors must have promoters, enhancers, polyA signals, selection cassettes, and expression augmenting elements (Jazayeri et al., 2018). The 9.567 kb pGFP/*DHFR*/*CYP21A2* plasmid vector will be constructed within T1 phage resistant MAX Efficiency™ DH5 α *E. coli* cells (ThermoFisher Scientific, Waltham, MA), using cDNA sequences that are integrated via restriction enzymes and DNA ligase. The hypothetical sequence for pGFP/*DHFR*/*CYP21A2* was designed using the current scientific literature on plasmid engineering, and it is outlined below:

P_{hCMV} human cytomegalovirus early promoter, met-loxP *GFP* green fluorescent protein minigene, poly(A)_{NSV40} SV40 poly A additional sequence,, *P_{SV40DE}* 72-bp repeated enhancer deleted from the SV40 early promoter, *DHFR* minigene in the form of a cosmid-derived 40 kb BglI fragment, (A)_{NBCH} Bovine growth hormone poly A additional sequence, *P_{SV40e}* SV40 early promoter, *CYP21A2* cytochrome P450 family 21 subfamily A member 2 minigene, (A)_{HTK} herpes simplex virus thymidylate kinase poly A additional sequence, *P_{SV40e}* SV40 early promoter, *PuroR* puromycin resistance gene, *LTR* long terminal repeat

gene, *pUC ori* pUC119 replication origin, *AmpR* ampicillin resistance gene, *P_{AmpR}* AmpR promoter, *LTR* long terminal repeat gene.



Figure 3. Graphical map of pGFP/DHFR/CYP21A2 plasmid vector showing all necessary restriction enzymes for gene integration. Created using GenSmart Design (Genscript Biotech)

6.1.3 Transfecting CHO-DG44 cells with pGFP/DHFR/CYP21A2 plasmid vector

Parental CHO-DG44 cells (ThermoFisher, Waltham, MA) will be initially suspended in a cell culture media consisting of CHO-S-SFMII with hypoxanthine and thymidine (ThermoFisher, Waltham, MA) supplemented with 5% premium fetal bovine serum (FBS) (R&D Systems, Minneapolis, MN). CHO-DG44 cells will then be centrifuged and washed in phosphate buffer saline (PBS); next, they will be resuspended in PBS on ice in a concentration of 5×10^6 cells/mL. pGFP/DHFR/CYP21A2 will be transfected into CHO-DG44 cells using the Gene Pulser MXcell™ electroporation system (Bio Rad, Hercules, CA), following all manufacturer's guidelines and procedures. Finally, cells that exhibit relatively high expression of GFP will be isolated via visual screening with a fluorescent microscope, and Southern blot analysis will then be used to determine the number of plasmid copies integrated into the cells' chromosomes.

6.1.4 Identifying transcriptionally active clones with fluorescence-activated cell sorter

Transfected cells with one or more copies of the GFP/DHFR/CYP21A2 transgene will then be transferred to tissue culture with CHO-S-SFMII media with 5% FBS, and they will be incubated at 37° Celsius for 24 hours. Then, transfectants will be centrifuged for 5 minutes at a speed of 1,000 rpm, collected, and resuspended in culture media consisting of HT-free CHO-S-SFMII (ThermoFisher, Waltham, MA), 5% FBS, and 5 nM MTX; cells will then be partitioned into 8-well plates with

approximately 6×10^3 cells in each well. After four weeks, the cells will be analyzed with a fluorescence-activated cell sorter (FACS). Separated cell colonies will be combined in a 6-centimeter dish, washed with PBS, treated with 0.025% trypsin in PBS for about 3 minutes at room temperature, and then centrifuged for 5 minutes at a speed of 1,000 rpm. The resultant cell pellet will then be suspended in 4° Celsius PBS in a concentration of approximately 10^6 cells/mL. Cells will be analyzed with the FACSymphony™ S6 (Becton Dickinson, Franklin Lakes, NJ) cell sorter to determine which clones have undergone gene amplification and therefore have greater GFP expression. Clones that exhibit the highest quantitative GFP expression via FACS analysis will be transferred to a 96-well plate with approximately 100 cells per well, suspended in a culture of HT-free CHO-S-SFMII supplemented with 5% FBS and 5 nM MTX until confluent. After three weeks, enzyme-linked immunosorbent assay (ELISA) will be used to assess 21-hydroxylase concentrations in the culture supernatants of each well. The cell colony with the greatest production level will then be transferred to a 250 mL spinner flask with EX-CELL 302 (Sigma Aldric, St. Louis, MO) media for 2 weeks, and 21-hydroxylase production will then be assessed with SDS-PAGE and FACS analysis.

6.1.5 MTX amplification

In the event that the first SDS-PAGE analysis of the initial cell line does not have adequate 21-hydroxylase production, further MTX amplification may be necessary. The 5 nM MTX gene-amplified cells will be treated with 25 nM MTX and evaluated via the same process outlined in the preceding section. If expression levels are still suboptimal, the process may need to be repeated once again with 50 nM MTX. Amplification with MTX concentrations of up to 500 nM may be necessary to achieve adequate 21-hydroxylase expression.

6.1.6 Collection of recombinant 21-hydroxylase enzymes

Once a stable, high-producing cell line has been identified, the cells will be suspended in EX-CELL 302 media inside 250 mL spinner flasks. Production will be continually monitored via SDS-PAGE analysis, and recombinant 21-hydroxylase enzymes will be purified, stored in Type 1 glass 5 mL vials, and refrigerated between 2° Celsius and 8° Celsius until they are prepared for infusion.

6.1.7 Labeling recombinant 21-hydroxylase enzymes via post-transcriptional modification

The 21-hydroxylase enzyme normally facilitates adrenal steroidogenesis from within the endoplasmic reticulum (ER) of adrenocortical cells. Thus, an essential consideration for this study is determining an effective method of labeling recombinant 21-hydroxylase enzymes for cell membrane endocytosis and transport to the ER. There is extremely little scientific literature regarding therapeutic proteins targeted to the ER, so the subsequent methodology will be based on the most up-to-date scientific knowledge of intracellular membrane

receptors, Sec61 channel polypeptide transport, and signal recognition particle (SRP) function; extensive *in vitro* testing will eventually be required to determine whether recombinant enzymes will successfully complete endocytosis through the cell membrane and be directed to the ER.

The Sec61 complex is a major component of the channel-forming translocon complex that mediates protein translocation across the endoplasmic reticulum (ER). The translocon complex is responsible for the post-translational ER import of nascent polypeptides. These nascent polypeptides are targeted to the ER via SRPs and ribosomes. The N-terminal signal sequence on the nascent polypeptide serves as the 'signal' that allows the ribosome-nascent chain complex (termed the RNC or cargo) to engage the SRP and, through interaction with the SRP receptor (SR), to be delivered to the vicinity of the Sec61 translocon within the ER membrane. The channel-forming subunit (Sec61 α) consists of 10 transmembrane helices and is arranged in two pseudo-symmetrical amino- and carboxy-terminal halves around a central constriction, which is sealed by a "pore ring." Once the RNC cargo is delivered in vicinity of the lateral gate, a hydrophobic patch in the cytosolic funnel of the engaged Sec61 complex serves as an interaction site for the incoming hydrophobic RNC signal peptide. This, in turn, displaces helix 2 of Sec61 α in order to destabilize the "lateral gate" and open the aqueous channel in the Sec61 complex for protein translocation. (Akopian et al., 2013; Favaloro et al., 2008; Lang et al., 2017).

In order to appropriately label recombinant 21-hydroxylase enzymes for endocytosis and ER uptake, two post-translational modifications will be made to the enzymes. Receptor-mediated endocytosis will be facilitated via the addition of a mannose-type oligosaccharide using N-linked glycosylation. Oligosaccharide transferase will be used to attach mannose-type oligosaccharides to an Asparagine side chain of recombinant 21-hydroxylase enzymes. Additionally, to facilitate SRP-ribosome complex recognition and ER uptake, a nascent polypeptide N-terminal signal sequence strand will be enzymatically added to 21-hydroxylase enzymes subsequent to their biosynthesis. This nascent polypeptide strand will signal SRP-ribosomal complexes to transport the recombinant 21-hydroxylase enzymes to the Sec61 complex, where they will then be translocated into the ER. It is not yet known whether the post-transcriptional addition of a mannose-type oligosaccharide or nascent polypeptide N-terminal signal sequence strand to recombinant 21-hydroxylase will interfere with the function of the enzyme. Western blot analysis of cells could elucidate whether 21-hydroxylase with these modifications are present in the ER and confirm the efficacy of this proposed strategy.

6.2 Subject selection

Due to the significant effect that 21-OHD has on development during adolescence, this proposed study will focus on pediatric patients in order to determine whether ERT could potentially improve height outcomes. Therefore, patients will only be selected for the study prior to epiphyseal fusion, which will be determined via radiography exams. Furthermore, in order to elucidate the effects of ERT on patients with varying degrees of enzymatic function, classic (salt-wasting and simple-virilizing) and non-classic 21-OHD patients will be tested. Roughly equal numbers of male and female patients are preferable, so that sex-specific

consequences of ERT can be determined. Ideally, the study cohort will consist of approximately 180 21-OHD patients between the ages 10 and 18 years: 60 classic salt-wasting, 60 classic simple-virilizing, and 60 non-classic. Identifying and enrolling this many participants within the set parameters will likely be difficult and pose a significant cost to investigators; however, a large sample size with maximum heterogeneity is essential for determining the efficacy of ERT for 21-OHD. Additionally, a control group of 180 of healthy individuals will be monitored simultaneously. All controls will be matched using anthropometric parameters, including age, sex, height, weight blood pressure, BMI, and smoking habits.

In order to maintain consistency between study participants, any patients undergoing treatment with gonadotropin releasing hormone analogs, aromatase inhibitors, or growth hormone will be excluded from the study. These medications would potentially alter results of the study and limit determinations as for the efficacy of ERT. Moreover, if ERT were to prove to be ineffective for maintaining proper hormone regulation, growth velocity, and bone aging, patients' final adult height could be compromised by stopping these medications in order to participate in the study.

6.3 Glucocorticoid and mineralocorticoid replacement therapy prior to study

All 21-OHD patients will be switched to a standard hydrocortisone regimen one week prior to the study, which will allow for an accurate baseline evaluation prior to the initiation of ERT. Patients will receive 12.5 mg of Cortef (Pfizer, New York, NY) at 08:00, 5 mg at 15:00, and 7.5 mg at 22:00 each day. If this dosage exceeds 25 mg/m²/day in any patients, they will have their dosage incrementally reduced by 7.5 mg (2.5 mg at each administration) until it no longer exceeds 25 mg/m²/day. Patients with salt-wasting 21-OHD will receive 100 µg of Florinef (Monarch Pharmaceuticals, Bristol, TN) at 08:00 each day.

6.4 Delivery method for ERT

Intravenous (IV) infusion of the recombinant 21-hydroxylase enzymes is the least invasive delivery method that will allow the enzyme to function properly in the adrenal glands. The adrenal glands have a high degree of vascularity and receive a significant blood supply, so there is little concern that infusion of recombinant 21-hydroxylase enzymes via the cephalic vein will fail to reach cells in the adrenal cortex. Moreover, the plasma half-life of recombinant enzymes typically ranges from 30 minutes to 120 minutes, though further testing specifically for the 21-hydroxylase enzyme may be required.

The first infusion will occur at 08:00, and all patients will have last taken hydrocortisone at 22:00 the previous day. Salt-wasting patients will have last taken fludrocortisone 24 hours prior.

IV solution will be formed by diluting concentrated recombinant 21-hydroxylase enzymes with 0.9% sodium chloride solution supplemented with 0.1% human Albumin. The dosage for ERT will be based on the bodyweight of the patient so that an adequate concentration of 21-hydroxylase is allowed to

diffuse within the adrenal glands. Patients in this study will be given 650 µg/kg body weight once per week via IV infusion. The initial rate of infusion will be 10 µg/kg/h, and if the medication is well tolerated, the infusion rate will be increased every 20 minutes to a maximum of 160 µg/kg/h. The actual volume of the total infusion and volumetric rate of infusion will vary based on the dosage of the individual patient and the concentration of recombinant 21-hydroxylase that is diluted with sodium chloride solution and Albumin.

Table 1. *Infusion rate of recombinant 21-hydroxylase*

Infusion Rate
10 µg/kg/h x 20 minutes
20 µg/kg/h x 20 minutes
40 µg/kg/h x 20 minutes
80 µg/kg/h x 20 minutes
160 µg/kg/h x remainder of infusion

The criteria for changing the infusion rate will be assessed prior to each incremental increase: heart rate, blood pressure, and signs of IARs will all be evaluated.

60 minutes prior to the first infusion, patients will be given antihistamines to reduce the risk of severe anaphylaxis or other IARs. If IARs are evident during the first infusion, then those patients will receive antihistamines 60 minutes prior to all future infusions. Due to the risk of adverse immunological responses to ERT, all infusions will be done in a hospital setting and home treatment will not be available at any point during the study.

7. Data to be Collected

7.1 Hormone analysis

The most immediate and predictable effect of ERT within 21-OHD patients should theoretically be a marked improvement in hormone regulation. Therefore, patients' hormone concentrations will be closely monitored for the duration of the study via frequent blood samples and subsequent serum and plasma analysis.

7.1.1 Serum and plasma assays

The following hormonal assay methods were defined in Verma et al., a pharmacokinetic and pharmacodynamic study. Serum concentrations of cortisol,

androstenedione, ACTH and insulin will be measured with a competitive chemiluminescence immunoassay. The cortisol assay has an analytical sensitivity of 5.5 nmol/l, with an intra-assay coefficient of variation (CV) of 3.2% and inter-assay CV of 3.6%. The androstenedione assay has a sensitivity of 1.0 nmol/l, average intra- and inter-assay CVs of 6.4% and 7.7% respectively. The ACTH assay has a sensitivity of 1.1 pmol/l, intra- and inter-assay CVs of 7.7% and 8.5% respectively. 17-OHP will be measured using liquid chromatography with tandem mass spectrometry, which has an analytical sensitivity of 0.12 nmol/l, inter-assay CV of 8.2% (Verma et al., 2010). Fasting plasma glucose will be measured using a glucose oxidase assay, which has an intra-assay CV of 5-7% and total analytical error < 8% (Sapin, 2003). Fasting serum insulin will be measured with ELISA.

Additionally, plasma aldosterone concentration, plasma renin activity (PRA), serum sodium, serum and potassium will be measured in all patients with the salt-wasting form of 21-OHD. Aldosterone and PRA will both be measured with liquid chromatography with tandem mass spectrometry. Serum sodium and potassium concentrations will be measured using ion-selective electrodes.

7.1.2 Blood sample procedures

A total of 26 blood samples will be taken during the 24-hour period subsequent to each ERT infusion. After receiving the infusion, patients will have blood drawn every 30 minutes for the first 2 hours, and then every 1 hour for the remainder of the 24-hour period. All samples will be analyzed for serum cortisol, androstenedione, 17-OHP, and ACTH. Blood samples in the salt-wasting 21-OHD group will be divided into separate test tubes, one treated with anticoagulants and one without anticoagulants. Salt-wasting patient's samples will be measured for plasma aldosterone concentration, plasma renin activity (PRA), serum sodium, serum and potassium.

7.2 Physical parameter measurements

All patients will have their height, weight, and bone age measured 24 hours prior to their first infusion. Height and weight will be recorded in centimeters and kilograms, respectively. Bone age will be assessed using hand and wrist radiographs according to the standards of the Greulich and Pyle method. Patients will have their height and weight measured on a weekly basis for the duration of the study. Bone age will be assessed every 2 months; patients will also have their bone age to chronological age (BA/CA) ratio calculated subsequent to each bone age radiograph.

7.3 ACTH stimulation test

All 21-OHD patients will undergo ACTH stimulation testing 72 hours prior to the first ERT infusion and 48 hours after the first ERT infusion. The tests will occur at 08:00, and patients will be asked to fast for at least 6 hours prior. Additionally, patients will not take their normal 08:00 dose of hydrocortisone on the day of their initial ACTH stimulation test (this will not apply for patients once ERT has commenced). Patients will have blood drawn prior to stimulation, then they will

receive an intramuscular injection of 250 µg cosyntropin (synthetic ACTH). After 60 minutes, blood will be drawn a second time. Both blood samples will be analyzed for 17-OHP concentration using high-pressure liquid chromatography with tandem mass spectrometry.

7.4 Quality of life factors

To assess relative quality of life across study groups, all 21-OHD patients and controls will be given a daily survey asking patients to rate several factors on a scale of 1-10: their energy levels during the morning (04:00 to 12:00) and afternoon (12:00 to 20:00), the quality of their sleep the previous night, and their overall physical and mental wellbeing. 21-OHD patients will begin answering this survey one week prior to the commencement of ERT, alongside all controls. A separate survey will be given to the parents/guardians of all 21-OHD patients and controls; the parent/guardian survey will have identical questions, though parents/guardians will be assessing the factors based on their observations rather than the children doing so themselves.

8. Hypothesized results

8.1 Statistical analysis

All data is presented as mean \pm standard error (SE). All analysis and calculations were done using Stata v17 (StatCorp, College Station, TX) and Vernier Graphical Analysis (Vernier, Beaverton, OR). Trapezoidal rule integration was used to calculate the area under the curve (AUC) of serum cortisol, ACTH, 17-OHP, and androstenedione. AUC was determined for the entire 24-hour observation period as well as three specific intervals: morning (04:00 to 12:00), afternoon (12:00 to 20:00), and nighttime (20:00 to 04:00). AUCs were calculated for 21-OHD patients receiving standard hydrocortisone (HC) treatment, 21-OHD patients during ERT, and controls. Additionally, the maximum concentration (C_{\max}) and the time at which it occurred (T_{\max}) was determined for cortisol, ACTH, 17-OHP, and androstenedione. Data was considered statistically significant with a p-value ≤ 0.05 , and p-values were determined using paired and unpaired *t*-tests, based on a 95% confidence interval and two-tailed hypothesis. P-values were computed for cortisol, ACTH, 17-OHP, and androstenedione, comparing all 3 study groups (patients during HC treatment, patients during ERT, and controls). All significant p-values are enumerated in Sections 8.2.1 - 8.2.4.

8.2 24-hour hormone profiles

The key factor for determining the clinical efficacy of ERT in the treatment of 21-OHD is the degree of prolonged hormone control achieved during therapy. Ideally, the increased bioavailability of the 21-hydroxylase enzyme within adrenocortical cells will allow for the endogenous, physiological synthesis of cortisol and aldosterone, without the buildup of precursor hormones. The 24-hour hormone profiles were hypothesized after consulting the most up-to-date scientific literature for normal physiology, current ERT results, and immunology:

Table 2. T_{max} and C_{max} of cortisol, ACTH, 17-OHP, and androstenedione of baseline HC therapy

	Cortisol	ACTH	17-OHP	Androstenedione
T_{max}	09:24 ± 18 min.	07:49 ± 14 min.	08:12 ± 13 min.	08:03 ± 8 min.
C_{max}	632.1 ± 38 nmol/L	19.1 ± 3 pmol/L	86.3 ± 11 nmol/L	7.5 ± 2 nmol/L

Table 3. T_{max} and C_{max} of cortisol, ACTH, 17-OHP, and androstenedione during ERT

	Cortisol	ACTH	17-OHP	Androstenedione
T_{max}	07:44 ± 12 min.	08:07 ± 12 min.	09:31 ± 9 min.	04:42 ± 17 min.
C_{max}	406.6 ± 22 nmol/L	8.5 ± 3 pmol/L	10.4 ± 2 nmol/L	3.2 ± 1 nmol/L

Table 4. T_{max} and C_{max} of cortisol, ACTH, 17-OHP, and androstenedione of controls

	Cortisol	ACTH	17-OHP	Androstenedione
T_{max}	07:58 ± 14 min.	08:11 ± 10 min.	09:34 ± 11 min.	04:51 ± 19 min.
C_{max}	576.2 ± 29 nmol/L	7.1 ± 2 pmol/L	4.2 ± 1 nmol/L	2.3 ± 1 nmol/L

Table 5. 24-hour AUC of hormones for participants during baseline HC therapy, participants during ERT, and controls

	Cortisol (h*nmol/L)	ACTH (h*pmol/L)	17-OHP (h*nmol/L)	Androstenedione (h*nmol/L)
HC	5802.4 ± 418	152.1 ± 11	615.3 ± 189	117.4 ± 28
ERT	3809.9 ± 372	93.2 ± 18	152.2 ± 42	46.4 ± 18
Controls	5463.1 ± 924	77.8 ± 21	60.9 ± 22	42.1 ± 22

Table 6. Morning (04:00 to 12:00) AUC of hormones for participants during baseline HC therapy, participants during ERT, and controls

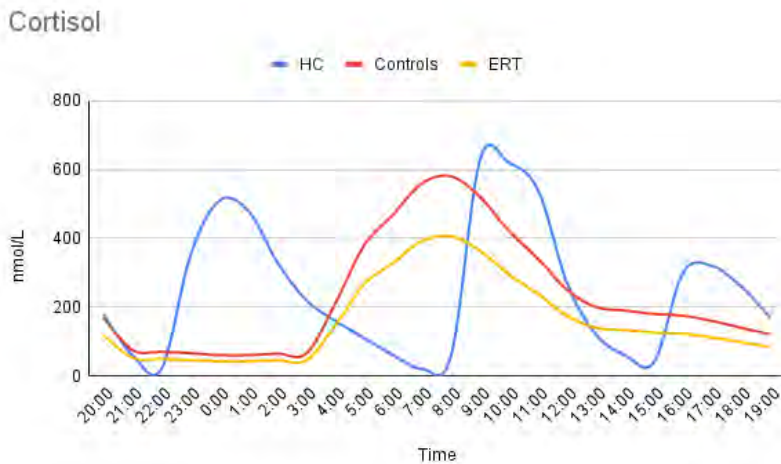
	Cortisol (h*nmol/L)	ACTH (h*pmol/L)	17-OHP (h*nmol/L)	Androstenedione (h*nmol/L)
HC	2255.4 ± 151	54.1 ± 9	279.9 ± 84	40.7 ± 14
ERT	2451.2 ± 118	41.2 ± 7	67.6 ± 24	17.1 ± 9
Controls	3501.5 ± 448	34.4 ± 7	27.1 ± 18	12.2 ± 8

Table 7. Afternoon (12:00 to 20:00) AUC of hormones for participants during baseline HC therapy, participants during ERT, and controls

	Cortisol (h*nmol/L)	ACTH (h*pmol/L)	17-OHP (h*nmol/L)	Androstenedione (h*nmol/L)
HC	1451.9 ± 139	71.5 ± 12	246.0 ± 99	43.3 ± 12
ERT	934.5 ± 178	30.9 ± 13	39.8 ± 11	16.2 ± 7
Controls	1355.2 ± 212	25.8 ± 10	15.9 ± 7	11.6 ± 4

Table 8. Nighttime (20:00 to 04:00) AUC of hormones for participants during baseline HC therapy, participants during ERT, and controls

	Cortisol (h*nmol/L)	ACTH (h*pmol/L)	17-OHP (h*nmol/L)	Androstenedione (h*nmol/L)
HC	2095.1 ± 249	26.5 ± 6	89.4 ± 44	33.4 ± 5
ERT	424.2 ± 143	21.1 ± 5	44.8 ± 22	18.3 ± 5
Controls	606.4 ± 298	17.6 ± 9	17.9 ± 9	13.1 ± 2

**Figure 4.** Predicted mean 24-hour serum cortisol concentrations of participants during baseline HC therapy, participants during ERT, and controls

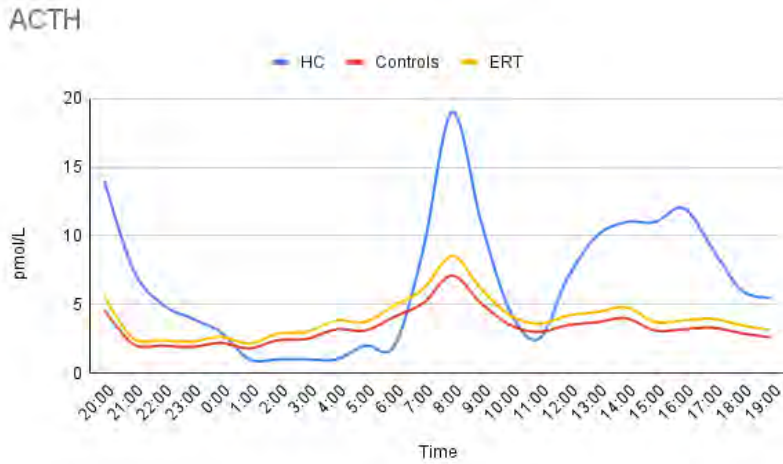


Figure 5. Predicted mean 24-hour serum ACTH concentrations of participants during baseline HC therapy, participants during ERT, and controls

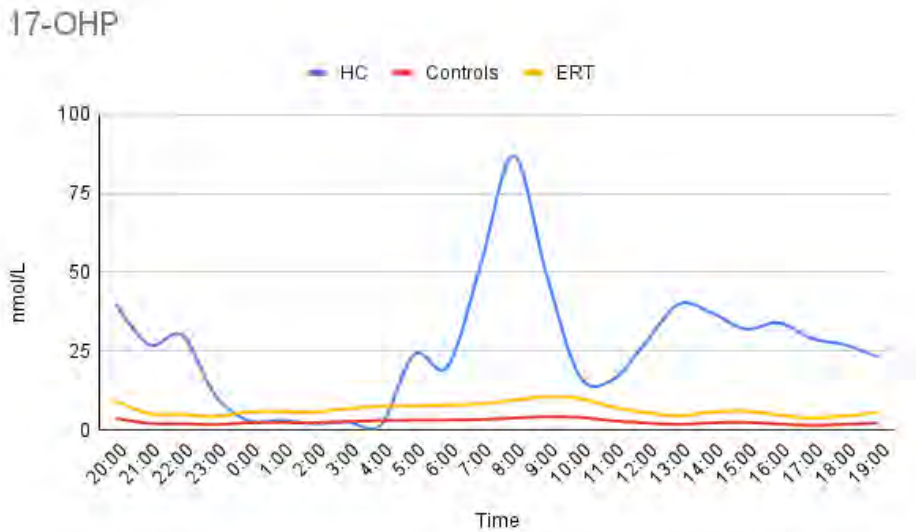


Figure 6. Predicted mean 24-hour serum 17-OHP concentrations of participants during baseline HC therapy, participants during ERT, and controls

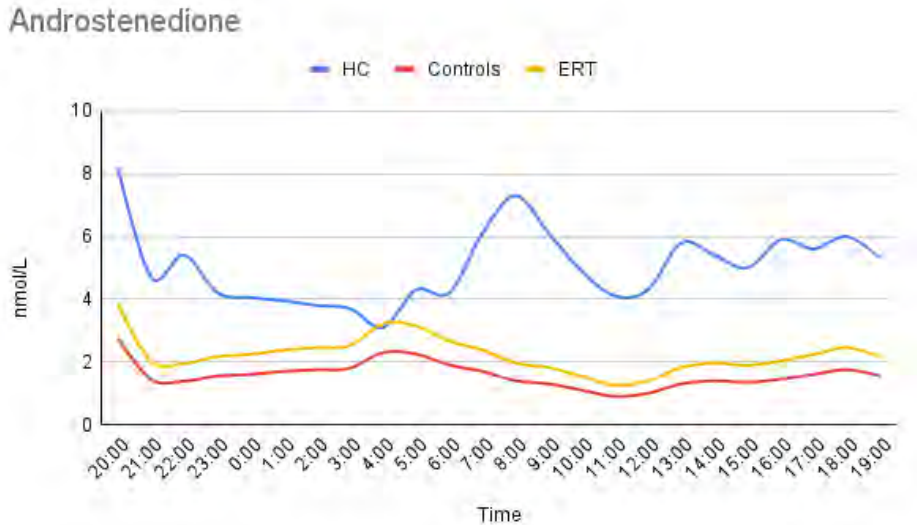


Figure 7. Predicted mean 24-hour serum androstenedione concentrations of participants during baseline HC therapy, participants during ERT, and controls

8.2.1 Cortisol

Analysis of hypothesized 24-hour cortisol concentrations indicated that total daily cortisol production did not significantly differ between patients receiving HC treatment and controls ($p = 0.7380$), nor did it significantly differ between patients receiving ERT and controls ($p = 0.1939$). This suggests that both HC treatment and ERT provide near-physiological levels of total cortisol production in a given 24-hour period. However, the intervals during which high and low cortisol concentrations occurred were markedly different between study groups. Notably, controls had significantly higher morning cortisol concentrations compared to patients receiving HC ($p = 0.0168$) and patients receiving ERT ($p = 0.0468$). Peak concentration (C_{\max}) occurred during the morning for all three study groups; however, C_{\max} was significantly lower in patients receiving ERT compared to controls ($p < 0.0001$) and patients receiving HC ($p < 0.0001$). Additionally, C_{\max} occurred significantly later in the morning for patients receiving HC compared to those receiving ERT ($p < 0.0001$) and controls ($p = 0.0003$). Afternoon cortisol concentrations, however, were not significantly different between controls and patients receiving HC ($p = 0.7029$) or ERT ($p = 0.2571$). The most pronounced difference in cortisol concentrations occurred during the nighttime period, during which patients treated with HC had

significantly higher concentrations than controls ($p = 0.0003$) and patients receiving ERT ($p < 0.0001$).

8.2.2 ACTH

Total 24-hour ACTH concentrations were significantly higher in patients receiving HC compared to those receiving ERT ($p = 0.0105$) and controls ($p = 0.0034$). Interestingly, the only specific time interval during which ACTH concentrations significantly differed between study groups was during the afternoon, when patients receiving HC had higher concentrations compared to patients receiving ERT ($p = 0.0435$) and controls ($p = 0.0069$). ACTH concentrations did not significantly differ across study groups during the morning and nighttime periods. C_{max} was also significantly higher in patients receiving HC compared to those receiving ERT ($p = 0.0249$) and controls ($p = 0.0017$); however, the time at which maximum concentration occurred did not significantly differ across study groups.

8.2.3 17-OHP

17-OHP concentrations over the total 24-hour period were significantly higher in patients receiving HC compared to those receiving ERT ($p = 0.0335$) and controls ($p = 0.0071$). The most notable difference in 17-OHP concentrations occurred during the morning, when patients receiving HC had significantly higher concentrations compared to patients receiving ERT ($p = 0.0302$) and controls ($p = 0.0065$). Afternoon 17-OHP was significantly higher in patients receiving HC compared to controls ($p = 0.0408$), though it was not significantly different when compared to patients receiving ERT ($p = 0.0769$). Nighttime 17-OHP concentrations did not significantly differ between the three study groups. Additionally, 17-OHP concentrations were not significantly different between controls and patients receiving ERT during any of the four measured intervals. C_{max} was significantly higher in patients receiving HC compared to those receiving ERT ($p < 0.0001$) and controls ($p < 0.0001$), and C_{max} occurred significantly later in patients receiving HC compared to those receiving ERT ($p < 0.0001$) and controls ($p < 0.0001$). Notably, C_{max} was also significantly higher for patients receiving ERT compared to controls ($p = 0.0111$); though, the time at which C_{max} occurred did not significantly differ ($p = 0.8328$).

8.2.4 Androstenedione

24-hour androstenedione concentrations exhibited the fewest significant differences out of the hormones measured, based on the hypothetical results of the proposed experiment. Total 24-hour androstenedione concentrations were higher, though not to a significant degree, in patients receiving HC and those receiving ERT ($p = 0.0659$) and controls ($p = 0.0689$). Morning androstenedione concentrations did not significantly differ across the 3 study groups. However, patients receiving HC did have significantly higher androstenedione than controls

during the afternoon ($p = 0.0244$) and the nighttime ($p = 0.003$) intervals. Furthermore, C_{\max} was significantly higher in patients receiving HC compared to controls ($p = 0.0401$), but it did not significantly differ between patients receiving HC and those receiving ERT ($p = 0.1090$). Though, C_{\max} occurred significantly later in the morning for patients receiving HC compared to controls ($p < 0.0001$) and patients receiving ERT ($p < 0.0001$).

8.3 Renin, aldosterone, sodium, and potassium

Plasma aldosterone concentration is predicted to be 2-7 ng/dL in salt-wasting 21-OHD patients receiving ERT. PRA is predicted to be 0.2-5.5 ng/mL/hour. The ratio of plasma aldosterone concentration to PRA should be < 30 ng/dL per ng/mL/hour in individual patients. Serum sodium should range from 130-150 mmol/L, and serum potassium should range from 3.4-4.8 mmol/L (Labcorp, 2021). Plasma aldosterone, PRA, sodium, and potassium concentrations should not significantly differ from age-related controls.

8.4 ACTH stimulation test

ACTH stimulation tests are widely considered the gold standard for diagnosing 21-OHD CAH, and results should indicate whether ERT has sufficiently increased the bioavailability of the 21-hydroxylase enzyme (Ospina et al., 2016). All patients receiving ERT are hypothesized to have stimulated 17-OHP concentrations of $< 1,200$ ng/dL, and baseline and stimulated 17-OHP concentrations should not be significantly different.

8.5 Linear height and weight velocity

Figure 7 and figure 8 illustrate normal height and weight velocities, respectively. During ERT, 21-OHD patients and controls are expected to have mean values of height and weight velocity between the 25th and 75th percentiles based on age, as illustrated in the Figure 7 and Figure 8.

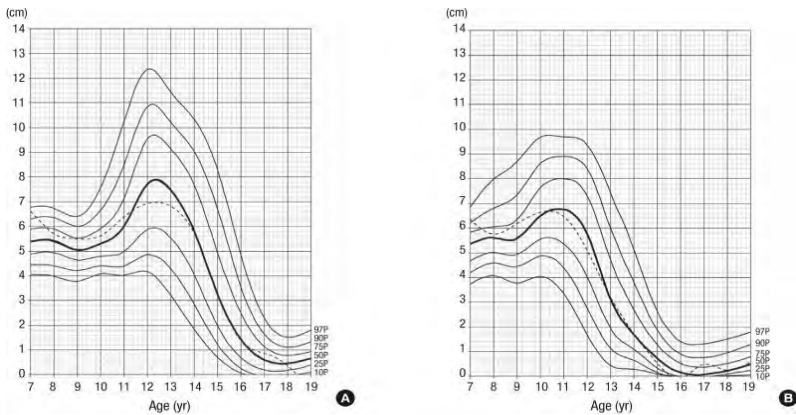


Figure 8. Growth velocity (cm/year) of pediatric males (left) and females (right) based on age (Chae et al., 2013).

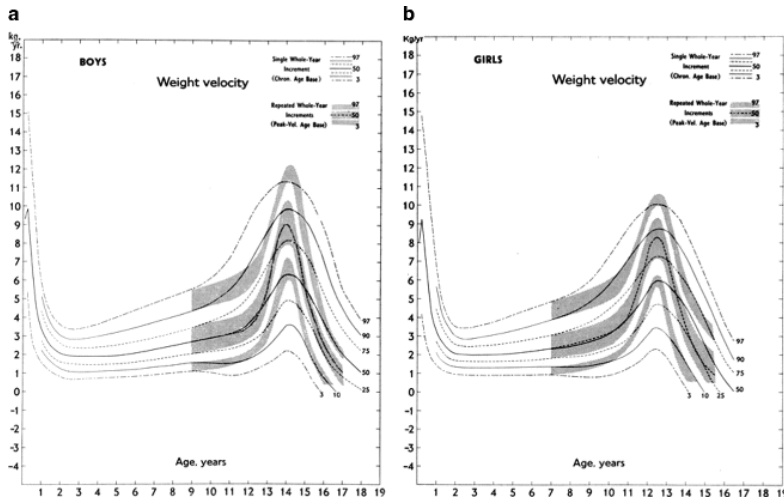


Figure 9. Weight velocity (kg/year) of pediatric males (left) and females (right) based on age (Bozzola et al., 2012).

8.6 Bone age and predicted final adult height

The ratio of patients' bone age to chronological age (BA/CA) will be calculated at regular intervals as outlined in Section 7.2, and it is predicted that BA/CA will decrease in patients receiving ERT therapy due to a marked reduction in the rate of epiphyseal maturation. This presumed decrease in the rate of skeletal maturation is based on the hypothetical decrease in 24-hour UAC of androgen precursors during ERT compared to baseline HC treatment; a decrease in subsequent androgen biosynthesis — testosterone, specifically — would then

curtail previously elevated estradiol production. Ideally, a patient's final adult height will be within one standard deviation of their mid-parental target height prior to epiphyseal fusion and growth cessation.

8.7 Metabolic profiles

8.7.1 Homeostatic model assessment for insulin resistance

In order to evaluate the impact of ERT on insulin sensitivity, the homeostatic model assessment for insulin resistance (HOMA-IR) modeling will be used to assess longitudinal changes in patients. HOMA-IR will be calculated using the following equation: fasting insulin (microIU/L) x fasting glucose (nmol/L) / 22.5. It is predicted that all controls will have HOMA-IR values between 0.5 and 1.4 (Salgado et al., 2010). Patients with 21-OHD will likely have varied HOMA-IR results prior to ERT, but it is hypothesized that HOMA-IR values should decrease significantly between the initial evaluation and 2-4 months after the commencement of ERT.

8.7.2 BMI

Average BMI of age-related controls will be compared to that of the 21-OHD patients at the beginning of the study. Mean BMI of 21-OHD patients will likely be significantly higher than controls, with significantly more patients considered overweight and obese. Theoretically, improved hormone profiles during ERT should lead to a decrease in BMI in overweight and obese patients. Therefore, patients are hypothesized to have a significantly lower BMI after 2-4 months of ERT treatment compared to baseline measurements.

8.8 Health-related quality of life outcomes

Based on the significant differences in 24-hour cortisol profiles across study groups, 21-OHD patients should theoretically report significantly higher energy levels in the afternoon during ERT compared to HC therapy due to the higher serum cortisol levels expected during that time interval. Furthermore, the quality of sleep should be significantly improved in 21-OHD patients during ERT compared to HC therapy considering the markedly lower cortisol concentrations during the nighttime period. Overall, energy levels should more closely resemble physiological circadian rhythm patterns during ERT compared to HC therapy. The results should, in theory, be identical between the patient-completed surveys and the parent-completed surveys; though, it is also possible that only patients would report a significant difference in quality of life since subtle changes in 24-hour hormone profiles may not be evident to parents.

9. Discussion

9.1 Rationale for proposed study

The purpose of this research proposal is to outline a potential therapeutic method for 21-OHD CAH that targets the underlying, enzymatic origin of the disorder. The current standard of care for 21-OHD is suboptimal, and modern therapeutics often introduce as many health problems as they alleviate. Investigators who oversaw a cross-sectional study of 199 adult 21-OHD patients concluded that “improvements in the clinical management of patients with CAH are required [...] and improving options for treatment in CAH is an important item on the research agenda” (Arlt et al., 2010), attesting to a need for novel therapeutic solutions. ERT is a potential method that would significantly reduce adverse health effects in patients and improve overall quality of life as it has already exhibited promising results for treatment of other enzymatic disorders. The facilitation of endogenous adrenal steroidogenesis via recombinant 21-hydroxylase enzyme supplementation should, in theory, be a vastly superior therapeutic method compared to conventional hormone replacement therapy (HRT); that is to say, ERT would likely restore patients’ normal circadian rhythm of hormone production, whereas typical strategies for HRT fail to replicate normal physiology. Therefore, it is thought that any patients who experience poor hormone control under conventional therapies would see markedly improved health outcomes after receiving ERT.

9.2 Interpretation of hypothesized results

The hypothetical results of this experiment were derived according to the presupposition that ERT — based on the proposed 650 µg/kg dosage — would restore 45-65% enzyme functionality, depending on the patient’s individual response to the recombinant enzymes. Corresponding, theoretical 24-hour hormone profiles were then formulated for 21-OHD patients who receive ERT, 21-OHD patients who receive conventional HC therapy, and healthy controls. The results of these 24-hour profiles illustrate markedly improved hormone control in patients who receive ERT compared to HC treatment.

Notably, there was no significant difference in total 24-hour cortisol levels between patients receiving ERT and healthy controls, suggesting that even partial enzyme functionality could allow for near-physiological cortisol biosynthesis. However, it should also be noted that patients receiving ERT have a significantly lower cortisol C_{max} than controls, possibly indicating that a higher dose of ERT may be necessary to ensure sufficient early-morning cortisol production. Patients receiving ERT and healthy controls also exhibited peak cortisol concentrations at ~ 07:50, whereas standard HC patients saw peak concentrations significantly later in the morning at ~ 09:30; this suggests that ERT would lead to improved circadian rhythm for cortisol production. Cortisol levels in ERT patients and controls also eclipsed those of HC patients during the nighttime and afternoon periods, indicating that ERT would allow 21-OHD patients to avoid instances of extremely low cortisol concentration typically seen in patients undergoing conventional treatment. The avoidance of the

aforementioned periods of very low cortisol concentration was further evidenced in the 24-hour profiles of ACTH, 17-OHP, and androstenedione. Specifically, 24-hour ACTH was significantly higher in patients receiving HC compared to those receiving ERT and controls, and this likely comes as a result of increased activation of the HPA axis due to periods of low cortisol concentration. Furthermore, maximum ACTH concentration occurred during the afternoon for all study groups, but it was markedly higher in patients receiving HC, which undoubtedly contributed to the subsequent rise in nocturnal androgen precursor hormones. Hence, total morning concentrations of 17-OHP were greater in patients receiving HC compared to ERT and controls, and 17-OHP C_{\max} was considerably higher in patients receiving HC and occurred earlier in the morning, at ~ 08:10, compared to patients receiving ERT and controls, for whom 17-OHP C_{\max} occurred at ~ 09:30. These instances of substantially elevated 17-OHP lead to increased biosynthesis of testosterone and DHT, which precipitates chronic hyperandrogenism in 21-OHD patients receiving HC. Androstenedione concentrations over the 24-hour period were higher in patients during HC therapy compared to patients during ERT and controls; however, the difference was not deemed significant ($p = 0.0659$ and 0.0689 , respectively). This near-significant result may indicate that differences in 24-hour androstenedione concentrations are a less pronounced indicator of differences in hormone control compared to ACTH and 17-OHP; though, it should still be considered a valid determinant of hormone regulation. Interestingly, afternoon and nighttime androstenedione concentrations were higher in patients during HC therapy compared to controls, though not when compared to patients receiving ERT. This result *prima facie* seems to indicate that ERT negligibly improves androstenedione concentrations compared to HC therapy; however, androstenedione concentrations were not significantly different between controls and patients receiving ERT during any of the four measured intervals, suggesting that adequate suppression of androgen precursors is achieved with ERT, whereas HC therapy fell short during the afternoon and night time intervals. Androstenedione C_{\max} was significantly higher in patients receiving HC versus controls, but there was no significant difference in C_{\max} between patients receiving HC compared to those receiving ERT, nor was there a difference between ERT and controls.

Overall, 24-hour hormone profiles for patients receiving ERT suggest that partial recovery of enzyme function facilitates near-physiological adrenal steroidogenesis. The most evident differences between treatment groups were the time intervals during which high and low concentrations of hormones occurred, and this is unquestionably caused by the inability for HC replacement therapy to replicate normal circadian rhythm. Notably, C_{\max} of cortisol, ACTH, 17-OHP, and androstenedione was higher in patients receiving HC, and these instances of glucocorticoid and androgen precursor excess are the cause of many of the health complications experienced by 21-OHD patients. Improved hormonal control via ERT would no doubt improve other health aspects, including height outcomes, BMI, insulin resistance, and accelerated bone aging. Moreover, benefits to these health aspects would likely be synonymous with improved quality of life for the majority of 21-OHD patients receiving ERT. Facilitating endogenous, adrenal steroidogenesis would also, theoretically,

greatly reduce instances of adrenal crisis caused by hypocortisolism, thus conferring a lower mortality rate amongst 21-OHD patients.

Based on the hypothetical results that patients receiving ERT would have significantly lower morning cortisol and consistently higher (though not significantly) androgens than controls, the dosage for ERT may need to be increased. Patients should be assessed on an individual basis for cortisol production and have their ERT doses adjusted accordingly. It is very likely that non-classic 21-OHD patients will require a lower dose than classic 21-OHD patients as they already maintain partial enzyme function. Further testing based on specific disease phenotype is necessary and linear regression should be used to chart a patient's normal enzyme activity and the ERT dose needed to replicate near-physiological cortisol production; theoretically, this model would produce an equation for proper dosage based on the severity of an individual's disorder.

9.3 Limitations and recommendations

The most obvious disadvantage for ERT compared to conventional oral steroid medications is the relative invasiveness of the treatment; receiving weekly infusions that could last up to 4 hours may actually reduce quality of life in patients when compared to 2-3 pills each day. Thus, ERT should be reserved for patients who experience poor hormone control when receiving the current standard of care, or patients for whom standard medications cause adverse health effects. ERT may also be more justifiable for pediatric patients as limiting chronic hyperandrogenism and glucocorticoid excess could substantially improve final adult height. Overall, the need for ERT versus conventional medication should eventually need to be considered based upon a patient-physician evaluation, where all health and quality of life impacts are considered.

Another limitation of this proposed experiment is the potential for an immunogenic response to recombinant enzymes. It is not yet known whether ADAs would form in 21-OHD patients receiving ERT, or if ADAs would have any impact on the efficacy or safety of the therapy. The more pressing concern, however, is the risk for acute anaphylaxis or other severe immunological responses to ERT. Pre-treatment with antihistamines should theoretically reduce the risk of such incidences. Furthermore, ERT should not be continued for patients who exhibit severe IARs at any point during treatment.

Finally, the cost to develop, produce, and deliver ERT is a substantial limitation to the potential therapeutic option. Based on costs for current ERT therapies for lysosomal storage disorders, the recombinant 21-hydroxylase enzymes could potentially cost hundreds of thousands of dollars per year for individual patients. Research funding may be better used for developing new therapeutics for disorders with a higher morbidity and severity. However, 21-OHD remains a valid subject for future studies due to the high frequency of the disorder and its many associated health complications.

Future studies for ERT in 21-OHD should also consider adult patients and compare sex-related results, focusing on possible resolutions of TARTs, PCOS, and cardiovascular disease.

10. Conclusion

Based on the current efficacy of ERT for the treatment of other enzymatic disorders and the hypothesized results of this experiment, it appears that ERT is a promising solution to vastly improve hormone control in 21-OHD patients. ERT would likely ameliorate several of the adverse health conditions associated with conventional treatments for 21-OHD, which would translate to improved quality of life. Future studies for ERT are necessary, and they should consider the implications for 21-OHD, as well as other enzymatic disorders. Addressing the underlying, enzymatic cause of the disorders via ERT is, in all probability, a better treatment solution when compared to the current standard of care, as facilitating endogenous enzyme function would theoretically restore normal physiology and remove the need for secondary medications.

Another, equally promising, method for treating genetic disorders, such as 21-OHD, is gene therapy, which could theoretically restore enzyme functionality by directly repairing the underlying genetic mutations; however, gene therapy has several technical and ethical limitations that have thus far prevented human trials. Therefore, it seems that ERT is the next logical step on the research agenda for 21-OHD CAH treatment.

References

- Ahmed, S., Soliman, A. T., Ramadan, M. A., Elawwa, A., Abugabal, A., Emam, M., & De Sanctis, V. (2019). Long-term prednisone versus hydrocortisone treatment in children with classic Congenital Adrenal Hyperplasia (CAH) and a brief review of the literature. *Acta bio-medica: Atenei Parmensis*, 90(3), 360–369. <https://doi.org/10.23750/abm.v90i3.8732>
- Akopian, D., Shen, K., Zhang, X., & Shan, S. O. (2013). Signal recognition particle: an essential protein-targeting machine. *Annual review of biochemistry*, 82, 693–721. <https://doi.org/10.1146/annurev-biochem-072711-1647322>
- Arlt, W., Willis, D. S., Wild, S. H., Krone, N., Doherty, E. J., Hahner, S., Han, T. S., Carroll, P. V., Conway, G. S., Rees, D. A., Stimson, R. H., Walker, B. R., Connell, J. M., & Ross, R. J. (2010). Health status of adults with congenital adrenal hyperplasia: A cohort study of 203 patients. *The Journal of Clinical Endocrinology & Metabolism*, 95(11), 5110–5121. <https://doi.org/10.1210/jc.2010-0917>
- Bozzola M., Meazza C. (2012) Growth Velocity Curves: What They Are and How to Use Them. *Handbook of Growth and Growth Monitoring in Health and Disease*. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1795-9_180
- Chae, H. W., Suh, I., Kwon, A. R., Kim, Y. J., Kim, Y. H., Kang, D. R., Kim, H. Y., Oh, S. M., Kim, H. C., Kim, D. H., & Kim, H.-S. (2013). Longitudinal standards for height and height velocity in Korean children and adolescents: The Kangwha cohort study. *Journal of Korean Medical Science*, 28(10), 1512. <https://doi.org/10.3346/jkms.2013.28.10.1512>

- Charmandari, E., Weise, M., Bornstein, S. R., Eisenhofer, G., Keil, M. F., Chrousos, G. P., Merke, D. P. (2002). Children with classic congenital adrenal hyperplasia have elevated serum leptin concentrations and insulin resistance: Potential clinical implications. *The Journal of Clinical Endocrinology & Metabolism*, 87(5), 2114–2120. <https://doi.org/10.1210/jcem.87.5.8456>
- Chrousos, G. P., Kino, T., Charmandari, E. (2009). Evaluation of the hypothalamic-pituitary-adrenal axis function in childhood and adolescence. *Neuro-immunomodulation*, 16(5), 272–283. <https://doi.org/10.1159/000216185>
- Clarke, L. A., Wraith, J. E., Beck, M., Kolodny, E. H., Pastores, G. M., Muenzer, J., Rapoport, D. M., Berger, K. I., Sidman, M., Kakkis, E. D., Cox, G. F. (2009). Long-term efficacy and safety of laronidase in the treatment of mucopolysaccharidosis i. *PEDIATRICS*, 123(1), 229–240. <https://doi.org/10.1542/peds.2007-3847>
- de Silva, K. S. H., Kanumakala, S., Brown, J. J., Jones, C. L., Warne, G. L. (2004). 24-Hour ambulatory blood pressure profile in patients with congenital adrenal Hyperplasia - a preliminary report. *Journal of Pediatric Endocrinology and Metabolism*, 17(8). <https://doi.org/10.1515/jpem.2004.17.8.1089>
- Falhammar, H., Frisén, L., Norrby, C., Hirschberg, A. L., Almqvist, C., Nordenskjöld, A., & Nordenström, A. (2014). Increased mortality in patients with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *The Journal of Clinical Endocrinology & Metabolism*, 99(12). <https://doi.org/10.1210/jc.2014-2957>
- Favaloro, V., Spasic, M., Schwappach, B., & Dobberstein, B. (2008). Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *Journal of Cell Science*, 121(11), 1832–1840. <https://doi.org/10.1242/jcs.020321>
- Find a test.* Labcorp. (2021). <https://www.labcorp.com/test-menu/search>.
- Goedegebuure, W. J., & Hokken-Koelega, A. (2019). Aromatase Inhibitor as Treatment for Severely Advanced Bone Age in Congenital Adrenal Hyperplasia: A Case Report. *Hormone research in Paediatrics*, 92(3), 209–213. <https://doi.org/10.1159/000501746>
- Gomes, L. G., Madureira, G., Mendonca, B. B., & Bachega, T. A. (2013). Mineralocorticoid replacement during infancy for salt wasting congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Clinics (São Paulo, Brazil)*, 68(2), 147–152. [https://doi.org/10.6061/clinics/2013\(02\)oa05](https://doi.org/10.6061/clinics/2013(02)oa05)
- Gorostidi, M., Sobrino, J., Segura, J., Sierra, C., de la Sierra, Á., Hernández del Rey, R., Vinyoles, E., Galcerán, J. M., López-Eady, M. D., Marín, R., Banegas, J. R., Sarría, A., Coca, A., Ruilope, L. M. (2007). Ambulatory blood pressure monitoring in hypertensive patients with high cardiovascular risk: a cross-sectional analysis of a 20 000-patient database in Spain. *Journal of Hypertension*, 25(5), 977–984. <https://doi.org/10.1097/hjh.0b013e32809874a2>
- Güven, A., Nurcan Cebeci, A., ; Hancili, S. (2014). Gonadotropin releasing hormone analog treatment in children with congenital adrenal hyperplasia complicated by central precocious puberty. *HORMONES*. <https://doi.org/10.14310/horm.2002.1555>

- Halper, A., Hooke, M. C., Gonzalez-Bolanos, M. T., Vanderburg, N., Tran, T. N., Torkelson, J., & Sarafoglou, K. (2017). Health-related quality of life in children with congenital adrenal hyperplasia. *Health and quality of life outcomes*, 15(1), 194. <https://doi.org/10.1186/s12955-017-0769-7>
- Han, T. S., Krone, N., Willis, D. S., Conway, G. S., Hahner, S., Rees, D. A., Stimson, R. H., Walker, B. R., Arlt, W., Ross, R. J., & United Kingdom congenital adrenal hyperplasia adult study executive (CaHASE) (2013). Quality of life in adults with congenital adrenal hyperplasia relates to glucocorticoid treatment, adiposity and insulin resistance. *European journal of endocrinology*, 168(6), 887–893. <https://doi.org/10.1530/EJE-13-0128>
- Han, T. S., Tijhuis, M. A., Lean, M. E., & Seidell, J. C. (1998). Quality of life in relation to overweight and body fat distribution. *American journal of public health*, 88(12), 1814–1820. <https://doi.org/10.2105/ajph.88.12.1814>
- Jazayeri, S. H., Amiri-Yekta, A., Bahrami, S., Gourabi, H., Sanati, M. H., ; Khorramizadeh, M. R. (2018). Vector and cell line engineering technologies toward recombinant protein expression in mammalian cell lines. *Applied Biochemistry and Biotechnology*, 185(4), 986–1003. <https://doi.org/10.1007/s12010-017-2689-8>
- Kamoun, M., Feki, M. M., Sfar, M. H., & Abid, M. (2013). Congenital adrenal hyperplasia: Treatment and outcomes. *Indian Journal of Endocrinology and Metabolism*, 17(Suppl 1), S14–S17. <https://doi.org/10.4103/2230-8210.119491>
- Kim, M. S., Dao-Tran, A., Davidowitz, E., Tseng, T., Gilsanz, V., Ryabets-Lienhard, A., Nguyen, E., & Geffner, M. E. (2016). Carotid intima-media thickness is associated with increased androgens in adolescents and young adults with classical congenital adrenal hyperplasia. *Hormone Research in Paediatrics*, 85(4), 242–249. <https://doi.org/10.1159/000444169>
- Lang, S., Pfeffer, S., Lee, P.-H., Cavalié, A., Helms, V., Förster, F., Zimmermann, R. (2017). An update on sec61 channel functions, mechanisms, and related diseases. *Frontiers in Physiology*, 8. <https://doi.org/10.3389/fphys.2017.00887>
- Lanoix, D., Plusquellec, P. (2013). Adverse effects of pollution on mental health: The stress hypothesis. *OA Evidence-Based Medicine*, 1(1). <https://doi.org/10.13172/2053-2636-1-1-572>
- Lorenz, M. W., Markus, H. S., Bots, M. L., Rosvall, M., Sitzer, M. (2007). Prediction of clinical cardiovascular events with carotid intima-media thickness. *Circulation*, 115(4), 459–467. <https://doi.org/10.1161/circulationaha.106.628875>
- M., K., S., I., Y., F., K., Y., T., S. (2002). Construction of engineered CHO strains for HIGH-LEVEL production of recombinant proteins. *Applied Microbiology and Biotechnology*, 60(4), 442–448. <https://doi.org/10.1007/s00253-002-1134-1>
- Marumudi, E., Khadgawat, R., Surana, V., Shabir, I., Joseph, A., & Ammini, A. C. (2013). Diagnosis and management of classical congenital adrenal hyperplasia. *Steroids*, 78(8), 741–746. <https://doi.org/10.1016/j.steroids.2013.04.007>

- New, M. I. (1993). Nonclassical congenital adrenal hyperplasia and the polycystic ovarian syndrome. *Annals of the New York Academy of Sciences*, 687(1 Intraovarian), 193–205. <https://doi.org/10.1111/j.1749-6632.1993.tb43866.x>
- Nimkarn, S., & New, M. I. (2007). Prenatal diagnosis and treatment of congenital adrenal hyperplasia owing to 21-hydroxylase deficiency. *Nature Clinical Practice Endocrinology Metabolism*, 3(5), 405–413. <https://doi.org/10.1038/ncpendmet0481>
- Nunberg, J. H., Kaufman, R. J., Schimke, R. T., Urlaub, G., Chasin, L. A. (1978). Amplified dihydrofolate REDUCTASE genes are localized to a homogeneously staining region of a single chromosome in a METHOTREXATE-RESISTANT Chinese hamster ovary cell line. *Proceedings of the National Academy of Sciences*, 75(11), 5553–5556. <https://doi.org/10.1073/pnas.75.11.5553>
- Ospina, N. S., Al Nofal, A., Bancos, I., Javed, A., Benkhadra, K., Kapoor, E., Lteif, A. N., Natt, N., & Murad, M. H. (2016). Acth stimulation tests for the diagnosis of adrenal insufficiency: Systematic review and meta-analysis. *The Journal of Clinical Endocrinology & Metabolism*, 101(2), 427–434. <https://doi.org/10.1210/jc.2015-1700>
- Oster, H., Challet, E., Ott, V., Arvat, E., de Kloet, E. R., Dijk, D.-J., Lightman, S., Vgontzas, A., & Van Cauter, E. (2016). The functional and clinical significance of the 24-hour rhythm of circulating glucocorticoids. *Endocrine Reviews*, 38(1), 3–45. <https://doi.org/10.1210/er.2015-1080>
- Padidela, R., & Hindmarsh, P. C. (2010). Mineralocorticoid deficiency and treatment in congenital adrenal hyperplasia. *International Journal of Pediatric Endocrinology*, 2010, 656925. <https://doi.org/10.1155/2010/656925>
- Papadakis, G., Kandaraki, E. A., Tseniklidi, E., Papalou, O., & Diamanti-Kandaraki, E. (2019). Polycystic Ovary Syndrome and NC-CAH: Distinct Characteristics and Common Findings. A Systematic Review. *Frontiers in Endocrinology*, 10, 388. <https://doi.org/10.3389/fendo.2019.00388>
- Placzek, M., Degitz, K., Schmidt, H., & Plewig, G. (1999). Acne fulminans in late-onset congenital adrenal hyperplasia. *The Lancet*, 354(9180), 739–740. [https://doi.org/10.1016/s0140-6736\(99\)01862-0](https://doi.org/10.1016/s0140-6736(99)01862-0)
- Quintos, J. B., Vogiatzi, M. G., Harbison, M. D., New, M. I. (2001). Growth hormone therapy alone or in combination with gonadotropin-releasing hormone analog therapy to improve the height deficit in children with congenital adrenal hyperplasia. *The Journal of Clinical Endocrinology Metabolism*, 86(4), 1511–1517. <https://doi.org/10.1210/jcem.86.4.7412>
- Rehman, Q., Lane, N. E. (2003). Effect of glucocorticoids on bone density. *Medical and Pediatric Oncology*, 41(3), 212–216. <https://doi.org/10.1002/mpo.10339>
- Roche, E. F., Charmandari, E., Dattani, M. T., Hindmarsh, P. C. (2003). Blood pressure in children and adolescents with congenital adrenal hyperplasia (21-hydroxylase deficiency): A preliminary report. *Clinical Endocrinology*, 58(5), 589–596. <https://doi.org/10.1046/j.1365-2265.2003.01757.x>

- Salgado, A. L., Carvalho, L. de, Oliveira, A. C., Santos, V. N., Vieira, J. G., & Parise, E. R. (2010). Insulin resistance INDEX (HOMA-IR) in the differentiation of patients with non-alcoholic fatty liver disease and healthy individuals. *Arquivos De Gastroenterologia*, 47(2), 165–169. <https://doi.org/10.1590/s0004-28032010000200009>
- Sapin R. (2003). Insulin assays: previously known and new analytical features. *Clinical Laboratory*, 49(3-4), 113–121.
- Sartorato, P., Zulian, E., Benedini, S., Mariniello, B., Schiavi, F., Bilora, F., Pozzan, G., Greggio, N., Pagnan, A., Mantero, F., Scaroni, C. (2007). Cardiovascular risk factors and ultrasound evaluation of intima-media thickness at common carotids, carotid bulbs, and femoral and abdominal aorta arteries in patients with classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *The Journal of Clinical Endocrinology Metabolism*, 92(3), 1015–1018. <https://doi.org/10.1210/jc.2006-1711>
- Savlov, D., O’Gorman, C., Urbach, S., & Pope, E. (2008). Case 1: Severe acne - not just skin deep. *Paediatrics & Child Health*, 13(6), 507–509. <https://doi.org/10.1093/pch/13.6.507>
- Sciannamblo, M., Russo, G., Cuccato, D., Chiumello, G., Mora, S. (2006). Reduced bone mineral density and increased BONE metabolism rate in young adult patients with 21-hydroxylase deficiency. *The Journal of Clinical Endocrinology Metabolism*, 91(11), 4453–4458. <https://doi.org/10.1210/jc.2005-2823>
- Speiser, P. W., Dupont, B., Rubinstein, P., Piazza, A., Kastelan, A., & New, M. I. (1985). High frequency of nonclassical steroid 21-hydroxylase deficiency. *American Journal of Human Genetics*, 37(4), 650–667.
- Speiser, P. W., Dupont, J., Zhu, D., Serrat, J., Buegeleisen, M., Tusie-Luna, M. T., Lesser, M., New, M. I., & White, P. C. (1992). Disease expression and molecular genotype in congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Journal of Clinical Investigation*, 90(2), 584–595. <https://doi.org/10.1172/jci115897>
- Trapp, C. M., & Oberfield, S. E. (2012). Recommendations for treatment of nonclassic congenital adrenal hyperplasia (NCCAH): an update. *Steroids*, 77(4), 342–346. <https://doi.org/10.1016/j.steroids.2011.12.009>
- Turcu, A. F., & Auchus, R. J. (2015). Adrenal steroidogenesis and congenital adrenal hyperplasia. *Endocrinology and Metabolism Clinics of North America*, 44(2), 275–296. <https://doi.org/10.1016/j.ecl.2015.02.002>
- U.S. National Library of Medicine. (2018, March 20). Gonadotropin releasing Hormone (gnrh) analogues. *LiverTox: Clinical and Research Information on Drug-Induced Liver Injury*. <https://www.ncbi.nlm.nih.gov/books/NBK547863/>.
- Valayannopoulos, V., Wijburg, F. A. (2011). Therapy for the mucopolysaccharidoses. *Rheumatology*, 50(suppl 5), v49–v59. <https://doi.org/10.1093/rheumatology/ker396>
- Verma, S., VanRyzin, C., Sinaii, N., Kim, M. S., Nieman, L. K., Ravindran, S., Calis, K. A., Arlt, W., Ross, R. J., Merke, D. P. (2010). A pharmacokinetic and pharma-codynamic study of delayed- and extended-release hydrocortisone (chronocort) vs. conventional hydrocortisone (cortef) in the treatment of congenital adrenal hyperplasia. *Clinical Endocrinology*, 72(4), 441–447. <https://doi.org/10.1111/j.1365-2265.2009.03636.x>

- Volkl, T. M. (2006). Obesity among children and adolescents with Classic congenital adrenal Hyperplasia due to 21-Hydroxylase Deficiency. *PEDIATRICS*, 117(1). <https://doi.org/10.1542/peds.2005-1005>
- Williams, R. M., Deeb, A., Ong, K. K., Bich, W., Murgatroyd, P. R., Hughes, I. A., Acerini, C. L. (2010). Insulin sensitivity and body composition in children with classical and nonclassical congenital adrenal hyperplasia. *Clinical Endocrinology*, 72(2), 155–160. <https://doi.org/10.1111/j.1365-2265.2009.03587.x>
- Wilson, R. C., Nimkarn, S., Dunic, M., Obeid, J., Azar, M. R., Najmabadi, H., Saffari, F., & New, M. I. (2007). Ethnic-specific distribution of mutations in 716 patients with congenital adrenal hyperplasia owing to 21-hydroxylase deficiency. *Molecular Genetics and Metabolism*, 90(4), 414–421. <https://doi.org/10.1016/j.ymgme.2006.12.005>



Along a Post-Colonial Spectrum: Considering Theory through an Analysis of Jordanian and Lebanese Identities

Isabella Lashley

Author Background: *Isabella Lashley grew up in the United States and currently attends The Chapin School in New York, New York in the United States. Her Pioneer research concentration was in the field of international relations and titled "Future Global Orders."*

Introduction

The relationships between colonies and colonizers during and after colonial rule are fundamental to understanding the history and theorizing the future of the world. There continue to be echoes of colonialism in the modern global order. Over 80% of the land in the world had been colonized before the formal and legal end of colonialism after World War II.¹ Still these colonial systems vary greatly, and I argue that more brutal colonial systems tend to result in a decided aversion of the colonized power towards the colonizing power, while less violent colonial systems have at times resulted in continued support and friendship between the individual countries and people.

The protectorate and mandate systems were relatively gentle variants of colonization when compared to more formalized and often more brutal relationships like those between Britain and India, and between Belgium and the Democratic Republic of the Congo. The protectorate system guaranteed one country general power over another, specifically in regard to international affairs.² The mandate system, uniquely created by the victorious powers in World War I, similarly granted governing power over another group to a victorious country, for the stated purpose of developing these areas, more accurately described as "Westernizing" areas like the Middle East.³

The Sykes Picot agreement was a 1916 treaty between France and Britain, authored by Mark Sykes and François Georges-Picot, to determine how the then-Ottoman empire would be split up, anticipating its defeat in World War I. The Sykes Picot agreement outlined the French Mandate of Syria and Lebanon, the British

¹ Saman Abdulqadir Dizayi, "Locating Identity Crisis in Postcolonial Theory: Fanon and Said," *Journal of Advanced Research in Social Sciences* 2, no. 1 (December 25, 2019): 79, <https://doi.org/10.33422/JARSS.2019.05.06>.

² Britannica, T. Editors of Encyclopaedia. "Protectorate." *Encyclopedia Brit-annica*, May 15, 2020. <https://www.britannica.com/topic/protectorate-international-relations>.

³ Mike Shuster, "The Middle East and the West: WWI and Beyond," NPR (NPR, August 20, 2004), <https://www.npr.org/templates/story/story.php?storyId=3860950>.

Mandate of Palestine and Transjordan, and the British Mandate for Mesopotamia.⁴ This paper will focus on French power in Lebanon and British power in Jordan, considering the evolving relationships between the colonized and colonizing countries by applying a post-colonial lens. While the colonial treatment of both Jordan and Lebanon were quite similar, the Jordanian and Lebanese attitudes towards their past colonizers are vastly different. So, how can this gap highlight the limits of modern post-colonial theory? Are current theories sufficient?

Post-Colonial Theory

Postcolonial theory considers how colonizing and colonized powers interact. It is a lens with which to look at the past, present, and future global orders, with room for the nuances of power, freedom, and self-determination. Colonial systems have truly shaped national identities and cultures around the world; thus, post-colonial theory seeks to understand these often-complex relationships and national consciousnesses. Post-colonial theory addresses fundamental ideas of how our world interacts and why, considering the trauma and exploitation of the past.

This paper will explore the diverging responses of colonized peoples after gaining freedom and consider how versions of a post-colonial identity can be defined through both the relationship between France and Lebanon and between Britain and Jordan, informed by the theories of Edward Said, Franz Fanon, and Homi Babha. All of these theorists propose different frameworks of what it means to understand a post-colonial country and world.

Edward Said's Orientalism

Said's *Orientalism* is a seminal work of post-colonial theory, the foundation for much of the post-colonial discourse to come. In his work, Said argues that there is a binary between the East and West, an inherent difference between the way that the Eastern world and Western world are perceived. Said proposes that the West is falsely perceived as a valorized and masculine entity meant to care for a feminine and delicate East. Simultaneously, Said explains that the East is also considered sly and cunning, wicked and sneaky, along with its supposed weakness. Said argues that this power differential influences every interaction between the East and West. It is this unfounded perception of the East as both exotic and inferior that creates a deep divide between the East and West, which led unfortunately to the dehumanization and exploitation of many places in the East by many places in the West, particularly the colonial and mandate systems.⁵

Frantz Fanon's The Wretched of the Earth

Frantz Fanon similarly argues that the colonizers dehumanize those who they colonize, reducing them to animals in their minds. He described that colonizers felt like those whom they colonized were so completely unlike them that they must even be a different species. Fanon, however, takes a radical approach to the post-colonial, arguing that the colonized power revolts against the colonizers angrily, knowing that

⁴ "The Avalon Project : The Sykes-Picot Agreement : 1916," The Avalon Project : The Sykes-Picot Agreement : 1916 (Lillian Goldman Law Library, 2008), https://avalon.law.yale.edu/20th_century/sykes.asp.

⁵ Edward W. Said, *Orientalism* (London: Penguin Books, 2019).

they are not as weak and uncivilized as the colonizers may think. Fanon also believes that those in the margins of society and who have the least, who he calls the *lumpenproletariat*, are able to first recognize the violence and exploitation of the colonial relationship. In *The Wretched of the Earth*, Fanon believes that there must be a passionate, and often violent, rebellion for a people to truly be rid of their colonizer's legacy.⁶

Homi Babha's Of Mimicry and Man: The Ambivalence of Colonial Discourse

Homi Babha's postcolonial theory tells us to expect a kind of mimicry after a colonial relationship, of the colonized group seeking to imitate the colonizers to achieve a sense of power and autonomy. Babha offers the idea of a hybrid, of combining the original core of the subaltern's identity with colonial advantages, like certain technologies or political systems. Similarly, Babha offers the idea of a kind of cultural camouflage. He proposes that the colonized group will adapt to the culture and customs of the colonizers, even after the end of a formal colonial relationship, not because they truly believe it is naturally better, but for survival and advancement. He believes that the subaltern, or the colonized group, has not fallen into Said's described orientalist trap of thinking themselves inferior to their colonizers in the West.⁷

I will now apply these theories to the cases of Jordan and Lebanon and how their people feel towards and connect or do not connect with their colonizers. In these cases, the post-colonial experiences of the peoples often align with only some or even none of these theories, truly highlighting that despite their merit, their application can never be accurate universally.

Case Study 1: Jordan and Britain

Historical Context

Jordan, in ancient times, was part of the biblical land of Canaan and then a part of the Roman empire. In more recent times, it was a territory within the Ottoman Empire, with major cities like Petra, Amman, and Aqaba garnering cultural and commercial relevance. Following World War I and the fall of the Ottoman Empire, the League of Nations granted Britain a protectorate over Transjordan, a territory that has since been redefined into many different nations. The Emirate of Transjordan, lasting from 1921 to 1946, became one of many British colonies.⁸

Britain, composed of the modern England, Northern Ireland, Scotland, and Wales, was, in ancient times, also a territory within the Roman empire. In more recent times, beginning in the 18th century, Britain became a great colonizing power. Along with some colonies like the thirteen that would become the United States, Britain, with London as its capital city, amassed the largest modern empire. During its peak in the early 20th century, British colonies made up 23% of the global population, with over 412 million people, and 24% of the Earth's land, with almost 4 million square

⁶ Fanon, Frantz. *The Wretched of the Earth* (New York, NY: Grove Press, 2011).

⁷ Bhabha, Homi. "Of Mimicry and Man: The Ambivalence of Colonial Discourse." October 28 (1984): 125-33. Accessed July 17, 2021. doi:10.2307/778467.

⁸ "Jordan," Britannica Academic (*Encyclopædia Britannica*), accessed July 17, 2021, <https://academic.eb.com/levels/collegiate/article/Jordan/106459>.

miles of land.⁹ Because of the size and power of the British empire, Britain became an international hegemonic power, with influence, language, and culture, spreading around the world.

The Emirate of Trans-jordan, a kind of mandate or protectorate system, ended after World War II, after a push against colonialism. However, the impacts of British colonialism have endured in Jordan since. Notably, Jordan's borders were not defined until gaining its independence, becoming distinguished from Transjordan, which included modern day Israel and Palestine. In fact, the borders of Jordan have been more heavily influenced by the establishment of Israel and the following Israeli-Palestinian conflict than by the British. Therefore, while Jordanian history will always have a post-colonial era, the borders and territories of Jordan have not been defined by British colonialism.

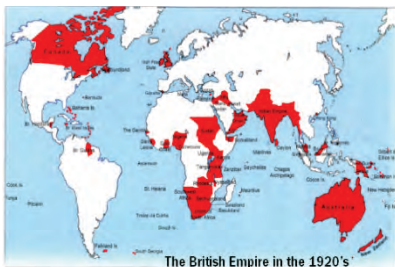


Figure 1. A map of the world in the 1920s with British colonies highlighted¹⁰

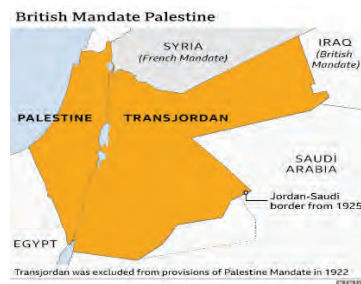


Figure 2. A map of the British mandates of Palestine and Trans-jordan¹¹



Figure 3. A map of modern-day Jordan and Its Borders¹²

⁹ "Britain," Britannica Academic (*Encyclopædia Britannica*), accessed July 17, 2021, <https://academic.eb.com/levels/collegiate/article/Britain/16506>.

¹⁰ Alcock, James. *Map of the British Empire in the 1920s* [Map]. Wikimedia, May 8 2006, https://upload.wikimedia.org/wikipedia/en/4/48/Map_of_the_British_Empire_in_the_1920%27s.png

¹¹ BBC, *British Mandate Palestine* [Map]. BBC News, September 16, <https://www.bbc.com/news/world-middle-east-54116567>

¹² Map of Jordan [Map]. *World Atlas*, <https://www.worldatlas.com/maps/jordan>

The Post-Colonial Pure Model

Eager to establish their sovereignty, the Jordanian people began to look for a symbol of their culture without colonial influence after gaining independence. The Bedouins fit this role perfectly: they served as an untouched model of the truly Jordanian, uncorrupted by colonial powers. The Bedouins, from the Arabic *بَدْوِي* or *بَدُو*, meaning “of the desert,” are Arab nomads who have roamed desert areas in the Middle East, including Jordan, since before the Common Era.¹³ The Bedouins and their so-called noble and original traditions and culture became an imagined pillar of Jordanian identity. Many Jordanians looked to the Bedouin example of their true, pure origin to heal the scar of the British colonial legacy.

Glubb Pasha

Although this is not exactly parallel, some British powers even exploited this Jordanian desire for freedom and identity during the protectorate, valorizing Bedouin style and traditions to disguise attempts to continue British influence in the emerging Jordanian military. Sir John Bagot Glubb, known as Glubb Pasha, was the Commander of the Transjordan’s Arab Legion from 1939 to 1956. He is credited with designing what would become the modern Jordanian military uniform with its trademark red and white keffiyeh.

Glubb took inspiration from traditional Bedouin keffiyeh used generally to protect from harsh Middle Eastern heat and wind. Glubb, however, sought to create a unique Jordanian identity through this keffiyeh, making the keffiyeh red and white rather than the Palestinian black and white, all the while ensuring that Jordanian identity would be inextricably tied to its colonial past. Glubb took further steps to weave British influence into the Jordanian military, creating the tradition of bagpipes at official Jordanian military events.¹⁴



Figure 4. *Bedouin falconer in traditional keffiyeh in Jordanian desert*¹⁵

¹³ Minkin, Shana. *The Arab Studies Journal* 9/10, no. 2/1 (2001): 142-45. Accessed July 17, 2021. <http://www.jstor.org/stable/27933817>.

¹⁴ Shryock, Andrew. *International Journal of Middle East Studies* 38, no. 3 (2006): 478-79. Accessed July 17, 2021. <http://www.jstor.org/stable/3879777>.

¹⁵ *Bedouin Falconer* [Photograph]. Michael Blackman, Ltd., Early 20th Century.



Figure 5. An illustration of Glubb Pasha in a keffiyeh and military uniform¹⁶



Figure 6. Bagpipes being played and red-and-white keffiyeh worn at official Jordanian military ceremony¹⁷

British Royals

Another reminder of the British mandate in Jordan is the British Jordanian royal family. King Hussein bin Talal (1935-1999) was the Jordanian king from 1952 to his death. The British mandate over Jordan only ended in 1946, so his reign was one of the first of a free Jordan. The first king of a free Jordan, Abdullah I, also had power during the mandate and the second king of a free Jordan, Talal, only reigned for one year before being forced to abdicate the throne.¹⁸ As such, King Hussein's choice of a bride and heir meant a lot to a post-colonial Jordan: one of their first sovereign queens and a new prince.



Figure 7. Princess Muna Al-Hussein with her sons, Abdullah II and Faisal in 1964¹⁹

<https://www.michaelbackmanltd.com/object/bedouin-falconry-stand/>

¹⁶ Lunt, James. Glubb Pasha, 1984, in *Glubb Pasha: A Biography of Lieutenant-General Sir John Bagot Glubb, Commander of the Arab Legion 1939-1956*, Bruegel (London: Harvill Press, 1984), Cover.

¹⁷ Kozlowski, Karol. *Jordanian Army Soldier playing his bagpipe* [photograph] Fine Art America, May 5, 2020. <https://fineartamerica.com/featured/jordanian-army-soldier-playing-his-bagpipe-karol-kozlowski.html>.

¹⁸ "Jordan," *Britannica Academic (Encyclopædia Britannica)*, accessed July 17, 2021, <https://academic.eb.com/levels/collegiate/article/Jordan/106459>.

¹⁹ Cozzi, Angelo. *Princess Muna with sons* [Photograph]. Wikimedia. March 24, 1964, https://commons.wikimedia.org/wiki/File:Princess_Muna_with_sons_1964.jpg#/media/File:Princess_Muna_with_sons_1964.jpg

In 1961, King Hussein married his second wife, Antoinette Gardiner, the daughter of a British army officer in Mandatory Palestine, from Suffolk, England. Jordan's new British queen, titled Muna Al-Hussein after marrying King Hussein, changed history. Similarly, when she gave birth to King Hussein's first male child, Abdullah II in 1962, the Jordanian royal family would be forever changed. Now, the heir to the Jordanian throne was half-British, and even descended from a British army officer who had worked within British mandate systems. While the Jordanian people were not hostile towards this British queen and half-British prince, they were instead consciously aware of their however partial "un-Jordanianness."

This fact became particularly apparent when Abdullah II ascended to the Jordanian throne. Abdullah had been educated mostly in the United States, attending Deerfield Academy in Massachusetts and Georgetown University in Washington DC and speaking English as his mother tongue. Therefore, he did not have native fluency in Arabic, particularly in formal Arabic, which would be essential to interact internationally with other Arabic-speaking leaders and nations. Many called him a Jordanian "Crown Prince with a British Accent."²⁰

The Jordanian people had reservations about this British affinity that was inextricable from Abdullah's identity, especially because of Jordan's colonial history. So, King Abdullah studied hard to achieve fluency in formal Arabic to complement his pre-existing fluency in casual spoken Jordanian Arabic, to finally be embraced by the Jordanian people.

Linguistic Independence and Preserving the Jordanian

Many Jordanians speak English and some are completely fluent in English. In fact, 45% of Jordanians can speak English proficiently.²¹ However, English is even more widely spoken among the educated and elite, since English is the primary language of advanced study in Jordan. Similarly, speaking English is more popular in Jordanian youth. In fact, some young Jordanians have named the composite of Arabic and English that they often use as the "Arabizi," a unique linguistic identity made up of Arab and British history and influence. Arabizi is a combination of the Arabic words for Arabic and English, Arabi and Inglizi. There have even been adaptations of the QWERTY keyboard to accommodate Arabic letters to be typed out, mostly in sentences of Arabizi, when both English and Arabic components, letters and phrases, are used to convey the message.²²

It is important to recognize that although the importance of speaking English is partly motivated by a globalizing world with an English-speaking United States as its hegemon, for Jordanians, speaking English has direct ties to their colonial past. Britain and its language claimed not only Jordanian political freedom but also Jordanian linguistic freedom. The English language has continued to be present and even important in modern Jordanian society, leaving echoes of colonialism in Jordan for the foreseeable future.

²⁰ "Crown Prince with an English Accent," *The Guardian* (Guardian News and Media, January 29, 1999), <https://www.theguardian.com/world/1999/jan/29/1>.

²¹ Majed Drbseh, "The Spread of English Language in Jordan Majed," *International Journal of Scientific and Research Publications* 3, no. 9 (September 2013): pp. 1-5.

²² "Arabizi, the Arabic Chat Language Changing the Way Young People Write & Speak," TeachMideast (TeachMideast, November 22, 2020), <https://teachmideast.org/articles/arabizi-the-arabic-chat-language-changing-the-way-young-people-write-speak/>.

Many Jordanian parents feel that early fluency in English is a way to ensure a brighter future for their children, and so encourage their children's English reading and writing fluency and just allow them to naturally learn to read and write in Arabic without any true support from their parents.²³ In line with Babha's theory of mimicry, these parents do not believe that English is superior to Arabic, just that English will better help their children to succeed in a world that values English so heavily.

However, some Jordanians are pushing back against the prevalence of English in Jordan. Some Arabic-speaking authors are writing children's books in Fusha, formal Arabic, to ensure that young children do not only read English and specifically British children's stories that are often used in Jordan to teach children English at a young age. This movement works towards ensuring that young children do not lose Arabic in favor of English. Authors like Taghreed Alnajjar along with publishing companies like Alsalwa write new stories and even translate English stories to Arabic to encourage children to build reading proficiency in formal Arabic as well. There are even YouTube channels dedicated to fostering a love of these books and of speaking Arabic for young children. One Jordanian YouTube account, Lina Samawi, has over 10,000 views on videos reading Arabic children's books aloud. Numbers like these show that many Jordanian families are trying to preserve the Arabic language by encouraging Arabic learning along with English.



Figure 8. Covers of Alsalwa publishing children's books (Source: Maktabee, 2021)

Applying Theory

I propose that for Jordan, since their national politics and finances have been stable since gaining independence, they are able to fully realize their own power and competency as a nation. Therefore, many Jordanian have a justifiable disdain for Britain and its power and resources and its exploitation of their country. Jordan maintained local stability with a clear line of succession politically, and thus did not have to forfeit independence and self-determination for the possibility of some semblance of structure and order from Britain. Jordan rejects this British support and influence because they are both resentful towards their colonial history and feel confident in their own abilities. Jordan rejects Said's theory of the Orient feeling less than the West, as Jordanians celebrate their independence and freedom, particularly from British influence. Jordan, however, does fit into Babha's theory of mimicry, with many valuing English greatly. Still, though, there are those fighting back against this mimicry in reinforcing the importance of Arabic. Finally, Jordan most closely fits into Fanon's theory, with the Jordanian people rebelling against all things English that remain in their society, albeit without Fanon's described violence. So, in the case of Jordan, some, while not all, post-colonial theory expects such outcomes as this.

²³ Drbseh, "The Spread of English Language in Jordan Majed," pp. 1-5.

Case Study 2: Lebanon and France

Historical Context

Lebanon, in ancient times, was the land of the Phoenicians. In more recent times, it was a territory within the Ottoman Empire, with Mount Lebanon enjoying slightly more autonomy than its counterparts in the coast and valley. It was at this time when European contact first became relevant to the Lebanese economy and history, with some European trading colonies established on the Lebanese coast. However, Lebanon was never formally a country until the French Mandate, which established the greater state of Lebanon within the French Mandate of Syria and Lebanon.²⁴ As such, Lebanon as a country and Lebanese identity is defined by French colonialism: it was the French Mandate that even designated the borders, defined the land, that would come to be Lebanon.

France, in ancient times, was the land of Gaul, a territory within the Roman empire. In more recent times, even during some politically and socially tumultuous years in the French Revolution or the following Reign of Terror, France established itself as a colonizing power. From as early as the 16th century, France began to conquer other parts of the world, mostly in the New World. In the following centuries, it maintained some of its colonies and established even more, becoming a global power of conquest, second only to the British Empire.²⁵ France had its colonial peak between 1919 and 1939, with 4,980,000 square miles of the globe under its control.²⁶

In 1923, the League of Nations officially declared Lebanon a mandate of France, a system and relationship ending roughly in 1946.²⁷ While the mandate that gave France power over Lebanon also gave France power over the land in modern-day Syria, the focus of this analysis will be the relationship between France and Lebanon. France treated Lebanon and its people with far less force and violence than it did some of its other colonies, like the Congo, because of the nature of a mandate, which is often far less exploitative and brutal than a formal colonial relationship.

The Lebanese proclaimed independence in 1943, and French troops slowly withdrew until 1946.²⁸ In the years following Lebanese independence, France and Lebanon have maintained a close relationship, with frequent diplomatic visits between countries, and Lebanese President Camille Chamoun in 1956 was even willing to risk an alliance with regional hegemon Egypt and its president Gamal Adel-

²⁴ "Lebanon," Britannica Academic (*Encyclopædia Britannica*), accessed July 17, 2021, <https://academic.eb.com/levels/collegiate/article/Lebanon/108463>.

²⁵ "France," Britannica Academic (*Encyclopædia Britannica*), accessed July 17, 2021, <https://academic.eb.com/levels/collegiate/article/France/110436>.

²⁶ "French Empire," French Empire - New World Encyclopedia (*New World Encyclopedia*), accessed July 17, 2021, https://www.newworldencyclopedia.org/entry/French_Empire#Second_French_colonial_empire.

²⁷ Reuters Staff, "TIMELINE: French-Lebanese Relations," *Reuters* (Thomson Reuters, November 18, 2007), <https://www.reuters.com/article/us-lebanon-france-relations/timeline-french-lebanese-relations-idUSL1842546320071118>.

²⁸ Bugh, G. Richard, Barnett, Richard David, Maksoud, Clovis F., Ochsenwald, William L., Kingston, Paul and Khalaf, Samir G.. "Lebanon." *Encyclopædia Britannica*, March 10, 2021. <https://www.britannica.com/place/Lebanon>.

Nassar to continue the French-Lebanese alliance.²⁹

Echoes of French colonialism have long rung out in Lebanon. Specifically, many Lebanese view France and all things French with a kind of adulation and prestige that alters what it means to be Lebanese. From language, to architecture, to immigration, many Lebanese have given undue praise and priority to France.

Language

French is widely spoken in Lebanon, with French taught equally as much in primary education as Arabic. More than 40% of the Lebanese population are Francophone.³⁰ In fact, many Lebanese can switch between French and Arabic, the Lebanese official language, even within the same sentence, with the response “kifek?” often being “ça va,” roughly translating to “how are you?” “I am fine.” Since Arabic and French are just as popular as each other in Lebanon, some Lebanese view this composite of Arabic and French as the “Lebanese mother tongue,” a uniquely Lebanese linguistic identity made up of Arab and French history and influence.³¹

In this way, France not only colonized the Lebanese people and nation but also their linguistic sovereignty. The French language lived on in Lebanon long past the official end of the mandate, a continuing reminder of French colonialism in Lebanon. Naturally, the French language became a part of Lebanese culture, blending into Arab and Middle Eastern identities to create a uniquely Lebanese affinity like Babha’s cultural hybrid.

Although French is not an official language of Lebanon, French-speakers are afforded more opportunities and prestige than those who do not speak French. The ability to speak French, unfortunately, is often a prerequisite to join the elite or politically powerful. Almost every recent Lebanese Prime Minister or President has been educated either in France or at a French-speaking Lycée in Lebanon. Tamaam Salam, a prior Lebanese Prime Minister and President, graduated from a Beirut Lycée. Michel Aoun, the current Lebanese president, attended Collège Notre Dame des Frères.

Still it is important to note that the French language and a French education do not outweigh the importance of also speaking Arabic fluently and formally. Although French is important in gaining entrance into the elite, Arabic is the dominant and official language of Lebanon, a language that every single Lebanese person can speak and understand well, regardless of class. As such, we can see that although the language of the colonizers is still popular in Lebanon, they have not abandoned their own language.

Architecture

Just like the hybrid of French and Arabic language in Lebanon, much of the architecture, particularly in Beirut, has both French and Ottoman-Arab influences. Beginning in the early 20th century, when France first gained influence over Lebanon,

²⁹ “Lebanon,” Britannica Academic (*Encyclopædia Britannica*), accessed July 17, 2021, <https://academic.eb.com/levels/collegiate/article/Lebanon/108463>.

³⁰ UNESCO, *Investing in Cultural Diversity and Intercultural Dialogue* (Paris: UNESCO, 2009).

³¹ Dalia Mortada, “Is Beirut the Codeswitching Capital of the World?,” *The World* (PRX, October 5, 2015), <https://www.pri.org/stories/2015-10-05/beirut-residents-switch-languages-practically-each-breath>.

many Lebanese buildings were adapted to reflect French architectural styles. These buildings are examples of early 20th century Beirut architecture and its French influences, like its arc-shaped windows and balconies. In fact, Beirut was once called the “Paris of the Middle East,” in reference to its French culture and architectural style.³²



Figure 9. *Early 20th century Beirut architecture, featuring pale yellow paint and balconies reminiscent of French colonial style (Source: Saliba, 1998)*

In 1975, the Lebanese Civil War began. It began as a result of building tensions between ethnic and religious populations in Lebanon, namely between Maronite Christians who had been favored during the French Mandate and allied with the Western World and Sunni Muslims who supported Pan-Arab ideologies and sided with Soviet-leaning countries. The war lasted 15 years until 1990, and killed approximately 120,000. The war destroyed much of the country and displaced millions of Lebanese, leading to millions fleeing the country, some even to France.³³



Figure 10. *Beirut architecture damaged after the Lebanese Civil War (Source: Tarek Adelsalam)*

When the Lebanese Civil War destroyed much of the early 20th century architecture, it also destroyed many of the physical influences of French architecture

³² Amani Sharif, “Why Beirut Was Once Called ‘The Paris Of The Middle East,’” Culture Trip (The Culture Trip, February 14, 2017), <https://theculturetrip.com/middle-east/lebanon/articles/why-beirut-was-once-called-the-paris-of-the-middle-east/>.

³³ “Lebanese Civil War,” Britannica Academic (*Encyclopædia Britannica*), accessed July 17, 2021, <https://academic.eb.com/levels/collegiate/article/Lebanese-Civil-War/472732>.

in Lebanon. There was a company formed, named the Lebanese company of the Development and Reconstruction of the Beirut Central District or SOLIDERE, commissioned to reconstruct the central district of Beirut, after the devastation of the war. This company sought to redefine the Beirut architectural identity and restore the historical identity of architecture in Beirut.³⁴

The company, however, reconstructed the buildings while maintaining the remnants of French colonialism. From pale yellow building facades to columned balconies, the Beirut identity itself was confused with a French colonial identity. The French influences in Beirut are so prevalent that they have become silent and invisible to those around them. SOLIDERE, in an attempt to create a solid Beirut architectural identity, instead chose to reintroduce and reconstruct the most tangible remnants of French colonial architecture. What it means to be Beirut is so wrapped up in what it means to be a French colony that they were confused with one another.



Figure 11. *Reconstructed Beirut buildings still with French colonial features (Source: Tarek Abdelsalam)*

SOLIDERE and other Lebanese had the opportunity to reclaim Lebanese independence and identity, apart from their French colonial past. They had the opportunity to rid themselves of any reminders of their lack of freedom and take control over their own nation. They had an opportunity to announce to the rest of Lebanon that Beirut could be more than just the Paris of the Middle East, to establish itself as more than a diluted copy, but as its own center of culture and commerce. Instead, they chose to reinforce the French colonial identity and remind the rest of Lebanon of its French colonial legacy, confirming the French colonial architectural vocabulary as now a part of Lebanon.

After the Beirut Explosion

Beirut, the capital city of Lebanon, was once considered the most Westernized city in the Middle East, due to its close ties with European forces, particularly France. However, despite its once prosperous ports and culturally rich institutions that blended the French and Arab, Beirut has since fallen on particularly hard times, financially and politically.

³⁴ Tarek Adbelsalam, "Reshaping the City Image: Impact of the French Colonial Architecture on the Contemporary Identity of Beirut Central District," Research Gate, January 2015, pp. 1-14.



Figure 12. *The aftermath of the Beirut explosion and the damage to the city*³⁵

On August 4th, 2020 at 6 pm local time, there was a giant explosion in a port in Beirut that killed hundreds, injured thousands, left hundreds of thousands homeless, and caused billions in damages. The blast was caused by the improper storage of ammonium nitrate, an explosive chemical material, inappropriately near residential areas and centers of commerce. The people of Beirut looked to the government for aid and were left disappointed, without sufficient intervention by government officials. Some even blamed the Lebanese Prime Minister Hassan Diab and other politicians for their negligence that may have led to the devastating explosion.³⁶

This lack of faith in their own government made some of the Lebanese people turn to their past colonizer: France. France, too, used this tragedy to further strengthen ties between Lebanon and France, and arguably to ensure that Lebanon stayed dependent in some way on French aid. In fact, unlike any other country, the French president Emmanuel Macron visited Beirut only two days after the explosion. During his visit, Macron spoke to locals in Beirut, who appeared to view him as someone to rely on, someone who they could trust over their own government. Many Lebanese already saw France as a golden land of opportunity and education offering political and financial stability, idolizing it despite their once colonial relationship; as such, after many Beirutis and Lebanese decided that they could not rely on their own government, France was their next choice.

Throughout his trip to Beirut, Macron made visits with generally unpopular Lebanese officials. He positioned himself and France as protectors of Lebanon even saying, “France will never let Lebanon go,” and “The heart of the French people still beats to the pulse of Beirut.” These almost blatant references to the past colonial relationship between Lebanon and France were also echoed and cheered on by some of the Lebanese people. There were even protests of the Lebanese government in front of the French Embassy in Beirut with chants like “Vive La France,” announcing to French ambassadors and to the world that many Lebanese had lost faith in their own

³⁵ Davies, Guy and Ibtissem Guenfoud, 1 American among 135 dead in massive Beirut explosion, officials say [Photograph]. ABC News, August 05, 2021, <https://abcnews.go.com/International/search-rescue-continues-beirut-wakes-effects-devastating-explosion/story?id=72185153>

³⁶ Nazih Osseiran and Isabel Coles, “Beirut Explosion: What Happened in Lebanon and Everything Else We Know,” *The Wall Street Journal* (Dow Jones & Company, December 10, 2020), <https://www.wsj.com/articles/beirut-explosion-what-happened-in-lebanon-and-everything-else-you-need-to-know-11596590426>.

government and had turned to France for political and financial rescue.³⁷ Some Beirutis went even further to display their allegiance to France, creating a petition to reinstate the French Mandate for Lebanon. More than 60,000 Lebanese signed this petition, making a public declaration of their desire to return to the semi-colonial relationship with France.³⁸

Applying Theory

The theory is silent on this situation and others like it: it is generally agreed upon that no country desires to be colonized, no matter how relatively less brutal their past colonial relationship was. Babha tells us to expect the Lebanese to view membership into a prescribed Frenchness as a necessary evil to gain prestige and power, not as inherently better. Babha views this pursuit as “profound and disturbing,” that much of postcolonial Lebanon has adopted this mimicry of France. Some Lebanese have gone further than imitating French culture and architecture, even requesting the return of the French mandate of Lebanon. Fanon’s theory is even more decidedly antithetical to this Lebanese petition, expecting that Lebanon emphatically hate France for having stripped away their control. So, how can 60,000 free Lebanese still want the return of the French Mandate system?

I advance that for Lebanon, their national politics and finances have been unstable and often filled with violence and corruption since gaining independence. Therefore, many Lebanese have undue admiration for France and its power and resources, solely because of the tumult in the Lebanese history without French colonialism, like the Lebanese Civil War. Those who signed the petition to return to a French Mandate craved colonial stability over local instability, willing to forfeit independence and self-determination for the possibility of some semblance of structure and order. In this perspective of France as a savior, Lebanon reenters the infantilization of colonialism, viewing themselves as less capable than France, their figurative colonial parent. Lebanon, now feeling inadequate and helpless, looks to France for the aid and guidance that a child might ask of a parent. Lebanon relishes this French support because France has the resources and expertise that some Lebanese feel their country lacks. Post-colonial theory does not expect such outcomes as these.

Conclusion

From Jordan to Lebanon, even in countries with such similar colonial experiences and such similar ethnic and linguistic identities, the post-colonial experience can be vastly different. Jordan benefitted from far more political stability than Lebanon enjoyed. Jordan’s borders were not so completely defined by their colonial history, unlike Lebanon. Jordanians have not gone as far as some Lebanese as to petition to reinstate the mandate system. These disparate post-colonial identities, despite

³⁷ Angela Charlton and Sarah Al-Deeb, “Is France Helping Lebanon, or Trying to Reconquer It?,” *The Washington Post* (WP Company, August 8, 2020), https://www.washingtonpost.com/world/middle_east/is-france-helping-lebanon-or-trying-to-reconquer-it/2020/08/08/3b582afe-d94b-11ea-a788-2ce86ce81129_story.html.

³⁸ Ines Djegham, “The Petition For a ‘French Mandate’ in Lebanon: 60,000 Signatures and More Questions,” *Human Rights Pulse* (Human Rights Pulse, October 19, 2020), <https://www.humanrightspulse.com/mastercontentblog/the-petition-for-a-french-mandate-in-lebanon-60000-signatures-and-more-questions>.

comparable colonial experiences, highlight just how subjective post-colonial experiences are. They cannot be predicted or anticipated: the same input can yield varying outcomes with regards to national identities, freedom, and a post-colonial world.

Fanon, Babha, and Said all expect a certain reaction, a certain sentiment from freed colonies. They all generally expect both a hatred for their past colonizers and a vague penchant towards continuing to speak the language and embrace the culture of their past colonizers. As many of the current iterations of post-colonial theory are not universally applicable to post-colonial experiences and identities, I advance that the post-colonial should be a spectrum, rather than a rigid and stagnant definition.

Still though, Jordan and Lebanon would not exist on the poles of this spectrum. They do indeed share some similarities in their post-colonial identities, namely the modern prevalence of the language of their colonizers. Accordingly, neither country could claim complete independence from their colonizers, which I propose would be one extreme of the spectrum. Similarly, neither country still has complete dependence on their colonizers, as Lebanon, despite its recent desire for French guidance, has also maintained a national identity separate from France. Jordan and France would lie between the poles and the center on opposite sides of this spectrum. They would be aligned with independence and dependence respectively, rather than completely identified with one or neither.

Current post-colonial theory is utilitarian, but insufficient. It does identify helpful, albeit general, patterns of post-colonial countries; however, it often forces wide and abstract concepts like self-determination, freedom, power, and exploitation into small and constricting categories. It does not allow for the grayness in which most international relationships live. Our current global order is evolving, with changing hegemony and new understandings of the implications of globalization. Along with this international evolution, the view on post-colonial relationships and national identity must be revolutionized to adapt alongside the continued salience of post-colonial relationships. Such comparative research opens further avenues of inquiry that may rely on a post-colonial lens.

Creating spaces in theory to assess different interactions and feelings allows us to understand not just the actions of nations but also their psyche, as a colonized or colonizing nation. At the core of international relations are individual countries and how they connect and interact, informed by the experiences and lessons of their past. Post-colonial theory is integral to understanding how our current world works and why. As such, it is paramount that post-colonial theory not be restrictive or definite or static; it must be able to move and reshape and adapt to fill the spaces that are now empty, applicable to mandates alongside colonies. The gray spaces that post-colonial theory should address can explain how to interpret our world and the nations within it, invaluable knowledge in an ever-evolving global order.

Works Cited

- Adbelsalam, Tarek. "Reshaping the City Image: Impact of the French Colonial Architecture on the Contemporary Identity of Beirut Central District." *Research Gate*, January 2015, 1–14.
- "Arabizi, the Arabic Chat Language Changing the Way Young People Write & Speak." TeachMideast. November 22, 2020. <https://teachmideast.org/articles/arabizi-the-arabic-chat-language-changing-the-way-young-people-write-speak/>.
- The Avalon Project: The Sykes-Picot Agreement: 1916. Lillian Goldman Law Library, 2008. https://avalon.law.yale.edu/20th_century/sykes.asp.
- Burney, Shehla. "Chapter Two: Edward Said and Postcolonial Theory: Disjunctured Identities and the Subaltern Voice." *Counterpoints* 417 (2012): 41-60. Accessed July 14, 2021. <http://www.jstor.org/stable/42981699>.
- Charlton, Angela, and Sarah Al-Deeb. "Is France Helping Lebanon, or Trying to Reconquer It?" *The Washington Post*. WP Company, August 8, 2020. https://www.washingtonpost.com/world/middle_east/is-france-helping-lebanon-or-trying-to-reconquer-it/2020/08/08/3b582afe-d94b-11ea-a788-2ce86ce81129_story.html.
- "Crown Prince with an English Accent." *The Guardian*. Guardian News and Media, January 29, 1999. <https://www.theguardian.com/world/1999/jan/29/1>.
- Djegham, Ines. "The Petition For a 'French Mandate' in Lebanon: 60,000 Signatures and More Questions." Human Rights Pulse. Human Rights Pulse, October 19, 2020. <https://www.humanrightspulse.com/mastercontentblog/the-petition-for-a-french-mandate-in-lebanon-60000-signatures-and-more-questions>.
- Drbseh, Majed. "The Spread of English Language in Jordan Majed." *International Journal of Scientific and Research Publications* 3, no. 9 (September 2013): 1–5.
- Fanon, Frantz. *The Wretched of the Earth*. New York, NY: Grove Press, 2011.
- "France." Britannica Academic. *Encyclopædia Britannica*. Accessed July 17, 2021. <https://academic.eb.com/levels/collegiate/article/France/110436>.
- "French Empire." French Empire - New World Encyclopedia. *New World Encyclopedia*. Accessed July 17, 2021. https://www.newworldencyclopedia.org/entry/French_Empire#Second_French_colonial_empire.
- "Lebanese Civil War." Britannica Academic. *Encyclopædia Britannica*. Accessed July 17, 2021. <https://academic.eb.com/levels/collegiate/article/Lebanese-Civil-War/472732>.
- "Lebanon." Britannica Academic. *Encyclopædia Britannica*. Accessed July 17, 2021. <https://academic.eb.com/levels/collegiate/article/Lebanon/108463>.
- Mortada, Dalia. "Is Beirut the Codeswitching Capital of the World?" *The World*. PRX, October 5, 2015. <https://www.pri.org/stories/2015-10-05/beirut-residents-switch-languages-practically-each-breath>.
- Osseiran, Nazih, and Isabel Coles. "Beirut Explosion: What Happened in Lebanon and Everything Else We Know." *The Wall Street Journal*. Dow Jones & Company, December 10, 2020. <https://www.wsj.com/articles/beirut-explosion-what-happened-in-lebanon-and-everything-else-you-need-to-know-11596590426>.

- Reuters Staff. "Timeline: French-Lebanese Relations." Reuters. Thomson Reuters, November 18, 2007. <https://www.reuters.com/article/us-lebanon-france-relations/timeline-french-lebanese-relations-idUSL1842546320071118>.
- Said, Edward W. *Orientalism*. London: Penguin Books, 2019.
- Sharif, Amani. "Why Beirut Was Once Called 'The Paris Of The Middle East'." Culture Trip. *The Culture Trip*, February 14, 2017. <https://theculturetrip.com/middle-east/lebanon/articles/why-beirut-was-once-called-the-paris-of-the-middle-east/>.
- Shuster, Mike. "The Middle East and the West: WWI and Beyond." NPR. NPR, August 20, 2004. <https://www.npr.org/templates/story/story.php?storyId=3860950>.
- UNESCO. *Investing in Cultural Diversity and Intercultural Dialogue*. Paris: UNESCO, 2009.



Effectiveness of Virtual Reality in Restoring Arm Function Following Stroke

Kenedy Quandt

Author Background: *Kenedy Quandt grew up in the United States and currently attends Sierra Canyon High School in Chatsworth, California in the United States. Her Pioneer research concentration was in the field of biology/neuroscience and titled “Treatment of Arm Dysfunction After a Stroke.”*

Abstract

Stroke is a common condition from which many survivors never fully recover; in fact, 26% of American stroke survivors need assistance to perform daily activities.⁴ Post-stroke patients experience various symptoms, such as spasticity, hemiparesis, muscle weakness, muscle tightness, and sensory loss. Virtual Reality is a relatively new intervention that aims to create a simulated virtual environment where patients can practice everyday skills and recover from post-stroke symptoms using activities and games. The objective of this review is to determine the effectiveness of Virtual Reality (VR) in improving post-stroke patients’ arm motor function and strength. To examine the effectiveness of VR, this review synthesizes the results of one meta-analysis, two systematic reviews, and eleven randomized control trials (RCTs). The RCTs are heterogenous in treatment intensity, inclusion/exclusion criteria, number of subjects, and methodology. Overall, results are mixed on the efficacy of VR on motor function, fine motor control, grip strength, and activities of daily living (ADLs).^{9-14, 17, 19, 20} Researchers have reported positive results in acute, subacute, and chronic stroke.^{10, 12, 13, 14, 17, 19, 20} However, due to conflicting evidence, more research is needed to determine the efficacy of VR intervention post-stroke.

Keywords

Stroke, Stroke Rehabilitation, Arm, Virtual Reality, Upper Extremity, VR, hemiparesis

1. Introduction

Cerebrovascular accident (CVA), commonly referred to as stroke, is a medical emergency that occurs when blood flow to brain neurons is disrupted, resulting in cell death. 87% of strokes are ischemic,¹ induced by a blood vessel blockage, while the other 13% of strokes are hemorrhagic, due to the rupture of a blood vessel. Both variants can be fatal; someone dies of stroke every four minutes in the United States.² Stroke remains a leading cause of long-term disability globally,² and many survivors cannot recover to the same level of function that they had prior to the stroke.³ In fact, 26% of American stroke survivors need assistance to perform daily activities.⁴ Consequently, stroke rehabilitation studies are crucial to uncover new information about the effects of stroke and to improve stroke survivors' quality of life.

After experiencing a stroke, survivors commonly experience spasticity, hemiparesis, muscle weakness, muscle tightness, sensory loss, and other forms of limb dysfunction because of brain damage. To help survivors with stroke-induced impairments recover and perform daily activities, therapists and researchers have searched for interventions that have the potential to improve movement in the arms and legs after stroke. Virtual reality (VR) therapy is being studied as a therapy option that integrates technology to provide "multi-sensory stimulation in real-time."^{5, page 1} Researchers suggest that cortical reorganization, a process that uses a stimulus to alter the brain's reaction to sensory input, is the underlying mechanism that makes VR rehabilitation work for patients post-stroke.⁷ Virtual reality utilizes a vast range of devices and computer programs, such as at-home gaming consoles, simple computer keyboard games, and at the most advanced level, VR headsets with sensory gloves. Immersive VR involves the use of a headset to make the user feel surrounded by the fabricated environment. In non-immersive VR, a user views a computer display from a distance. VR devices are becoming more popular because programs can be implemented at home and modified to match patients' needs. Through daily drills, training, games, and other methods, VR intervention immerses patients in a fabricated environment, either realistic or fictional, based on their desired outcomes and current capacity. Patients view VR intervention as motivating and approachable, which might mean they will spend more time in therapy and regain additional motor function and strength.⁵ Researchers hope that increasing the use of VR technology alone or alongside other treatments can improve motor learning by maximizing the repetition and intensity of tasks. VR devices not only provide long-term progress information to therapists but also can directly give auditory, sensory, and visual feedback to the user without requiring the presence of a therapist. VR intervention after a stroke gives users the opportunity to practice everyday skills that cannot be rehearsed in a hospital environment.⁶ Preliminary studies indicate that VR therapy potentially improves motor function, strength, arm function, and daily living capabilities in patients with chronic post-stroke hemiparesis. However, there is conflicting evidence suggesting that VR alone is not more effective than conventional therapy measures.⁵

No life-threatening risks have been reported from this treatment because the activities are self-paced and controlled by the user.⁶ VR is appropriate for stroke survivors who have no visual impairment and do not experience nausea and vertigo from a head-mounted display.⁵

This paper will provide a scoping review on the effectiveness of VR after stroke. Foremost, a thorough review will be conducted to examine if current research supports the continued use of virtual technology with stroke patients.

2. Methodology

This scoping review examines two systematic reviews, one meta-analysis, and eleven randomized control trials (RCTs) to assess the effectiveness of VR. Relevant studies for this scoping review were primarily selected using The National Center for Biotechnology Information, Cochrane Reviews, PubMed, and Embase. In addition to these studies on VR, one scoping review, one meta-analysis, and three RCTs on mirror therapy were used to evaluate a comparison treatment. Only papers with publication dates after 2005 were considered. The papers selected were limited to peer reviewed publications. Exclusionary factors for RCT studies included sample sizes below 12 subjects, no standardized tests for outcome measures, or high risk of author bias. Figure 1 contains a PRISMA diagram for VR study selection.

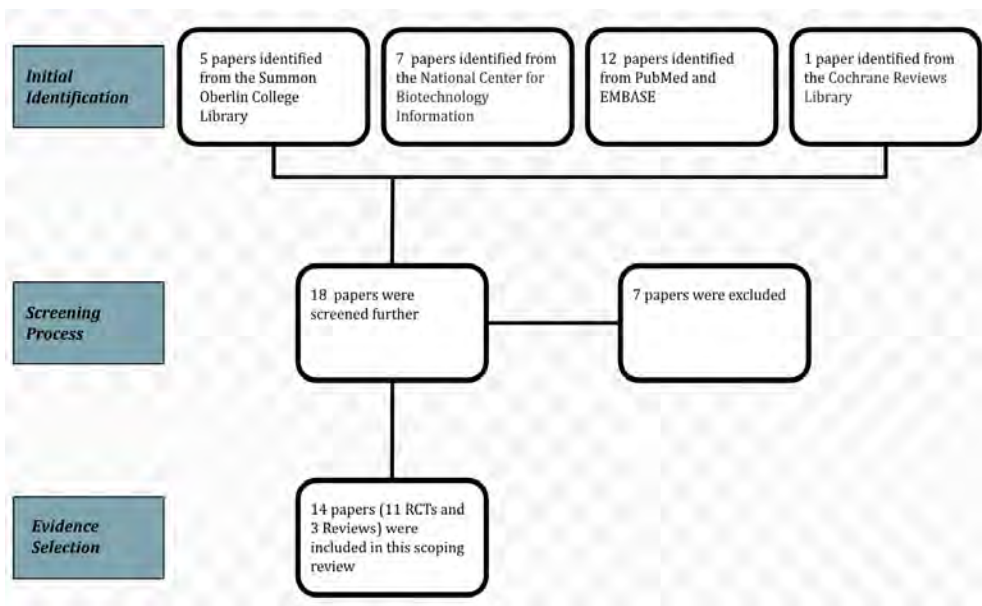


Figure 1. A Summary of the Search Strategy and Evidence Selection Process

3. Literature Review

In this section, the effectiveness of VR in stroke rehabilitation of the upper extremity will be evaluated by three criteria: technology, time post-stroke of participants, and outcome measures. The intensity and duration of VR intervention in the studies as well as whether virtual reality is used independently

or in conjunction with other treatments will also be considered. Appendix 1 provides a summary of evidence with detailed information on each study.

3.1. Virtual Reality Effectiveness by Construct

The reviewed studies evaluated the effectiveness of VR on hand and arm motor function using a wide variety of tests. Ten randomized control trials^{8-14, 17, 19, 20} used a total of twenty-four different outcome measures to examine VR effects. Appendix 2 provides a list of all the outcome measures used and the constructs examined. The three most common outcome measures utilized were the Fugl Meyer Upper Extremity (FMUE) test, the Action Research Arm test (ARAT), and the Functional Independence Measure (FIM). The type of test used is important to consider because each one evaluates different aspects of the arm.

The FMUE test was used by seven trials.^{9, 10, 12-14, 19, 20} The FMUE test measures motor function or movement in an affected limb following stroke. Six studies^{9, 12-14, 19, 20} reported statistically significant improvement pre-post for the VR group, five studies^{12-14, 19, 20} reported significant between group effects favoring the VR group, and one study⁹ reported non-significant between group differences. These results show that VR technology can be an effective component of stroke rehabilitation, and this technology should be studied further.

The ARAT was utilized in four studies.^{8, 10, 11, 19} The ARAT assesses changes in arm function by testing pinch, grip, grasp, and gross arm movement and a patient's capacity for holding objects of different sizes and weights. Three trials^{8, 11, 19} showed statistically significant improvement pre-post for the VR group, but only one study¹⁹ reported statistically significant effects on the ARAT in favor of the VR group compared to the control group. Based on this evidence, VR alone has not been found to be more effective than an equal time dose of conventional therapy in helping patients regain the ability to grasp different objects.

Three studies^{14, 17, 19} used FIM as an outcome measure. The FIM examines independence in daily activities. All studies^{14, 17, 19} demonstrated statistically significant pre-post improvement in VR groups. Further, all three studies^{14, 17, 19} reported significant between group comparisons that favored the VR group on the FIM. These results suggest that VR therapy helps patients decrease dependence in daily life.

3.2. Virtual Reality Effectiveness by Stroke Acuity

The recovery period, or time since stroke onset, of subjects is typically described as acute (less than 1 month), subacute (1-6 months), and chronic (over 6 months).²¹ Of the seven RCTs^{8-10, 12-14, 20} that examined non-immersive VR, three studies^{9, 13, 20} enrolled chronic subjects. While two of these studies^{9, 13} reported that subjects in both groups made significant improvements in outcome measures, one study²⁰ reported that only the VR group showed significant pre-post improvement in the Manual Muscle Test, Modified Ashworth scale, Box and Block test, 9-Hole Peg test, and the FMUE. These within group comparisons suggest that differences in group outcomes are affected by the type of VR technology used, the intensity of exercise, and the duration of treatment. Two studies^{13, 20} reported significant between group differences favoring the VR group.

In contrast, one study reported non-significant between group differences in the FMUE, Motor Proficiency Speed, Box and Block, and the Tapper tests.⁹

Across the seven RCTs^{8-10, 12-14, 20} 8, 10, 12, 14 that examined non-immersive VR, four^{8, 10, 12, 14} enrolled subjects across stroke acuity. Two of those studies^{10, 12} occurred within a month of stroke. One of these acute studies¹⁰ showed that the high intensity VR group had greater increases in the FMUE test and ARAT compared to a low intensity group; however, no analysis of statistical significance was performed. Nineteen subjects with acute stroke were recruited for another study¹² that reported the VR group showed significant improvements and faster improvements in paretic arm speed on the FMUE and on the Chedoke Arm and Hand Activity Inventory compared to a control group. One non-immersive study⁸ included acute and subacute subjects. There were no significant differences on the ARAT or Motor Activity Log between VR and control groups.⁸ Eighty patients in all three stages of recovery were recruited in another trial¹⁴ that compared VR plus standard care with a dose-matched control group. A statistically significant between group comparison for the FMUE test and the FIM was reported. However, there was no statistically significant between group difference in spasticity.¹⁴ The study¹⁴ also concluded that the use of reinforced feedback in a virtual environment plus standard care was statistically more effective than the dose-matched standard care for both ischemic and hemorrhagic stroke subjects. The inclusion of subjects across acuity is beneficial as it may help target intervention to patient characteristics.

Of the three RCTs that examined immersive VR included in this review^{11, 17, 19}, one study¹⁹ enrolled chronic subjects. In this study, the VR group had significantly more improvement on tests of self-care and motor function.¹⁹ This between group comparison exemplifies that VR intervention may be more beneficial to help patients in the chronic phase gain functional independence than conventional therapy alone.

Two immersive VR studies^{11, 17} enrolled subjects across the continuum of recovery. Both studies^{11, 17} reported no significant between group differences in motor function, arm movement, and spasticity. Since these studies mixed patients in different phases of recovery and performed no correlational statistics, they were unable to determine which acuity groups saw the least and most pre-post changes and if acuity was a factor that contributed to the lack of significant results.

Regardless of the type of VR technology used, none of the reviewed studies examined the optimal acuity to implement VR post-stroke. Since positive results can be seen in all stages of recovery and across outcomes, more testing is needed to determine if there is an optimal time post-stroke to deliver VR intervention.^{10, 12, 13, 14, 17, 19, 20}

3.3. Virtual Reality Alone vs. in Combination with other Treatments

When examining VR therapy, it is important to note if the experimental group is receiving VR alone or paired with another treatment because this may impact interpretation and generalizability of the study results. Four trials^{10, 12, 14, 17} examined VR therapy in conjunction with standard care, and three of these studies^{12, 14, 17} had sample sizes large enough to perform statistical analyses. Two studies^{12, 14} reported significant pre-post improvements in the FMUE test for the

VR group and reported that VR therapy paired with standard care was significantly more effective than the control group. The third study¹⁷ reported significantly more improvement in the VR + standard care group on the FIM but no significant between group difference on the Brunnstrom Stages test. These results suggest that incorporating VR therapy into an existing standard care regime may be beneficial for patients.

Many studies examined isolated VR therapy. Six trials^{8, 9, 11, 13, 19, 20} compared an isolated VR group to standard care. Five of those studies^{9, 11, 13, 19, 20} reported significant pre-post improvements in the outcome measures for the VR group, and three studies^{13, 19, 20} reported that isolated VR therapy was significantly more effective than standard care. Three studies^{8, 9, 11} reported non-significant between group differences, meaning the use of isolated VR was not more effective than conventional therapy measures. These conflicting findings show that more studies are needed to compare isolated VR to standard care to determine VR's value for patients in recovery.

3.4. Location of Virtual Reality Intervention

Another important factor to consider is the location of treatment. One study⁸ that took place at subjects' homes had experimental subjects use Nintendo Wii Sports alone for up to 45 minutes/day for six weeks while a control group performed arm exercises in a seated position for the same amount of time. Since this treatment was performed by patients at home without supervision by a therapist, subjects did not have to pay for time in the clinic or for transportation. No significant difference between the VR group and control group was reported on the ARAT or Motor Activity Log.⁸ Despite this, at home systems potentially have the capacity to advance activities based on performance and to increase training hours for patients. More research should be done to determine the effectiveness of VR therapy at home because it can be more cost effective and saves patients' time.

3.5. Non-immersive Virtual Reality Effectiveness

Seven RCTs evaluated non-immersive VR.^{8-10, 12-14, 20} Four of these studies^{12-14, 20} compared non-immersive VR to standard care and reported statistically significant improvements in strength, function, and spasticity favoring the VR group. However, two studies^{8,9} reported no statistically significant differences between groups in motor function, fine motor control, strength, and movement. These results demonstrate that non-immersive VR may be beneficial but requires more research to determine its efficacy in stroke rehabilitation.

Although researchers in one study¹⁰ were unable to perform statistical analysis because of the small sample size, this study remains relevant because descriptive statistics demonstrate greater pre-post score changes on the FMUE and ARAT for subjects in the high-intensity VR group compared to a low-intensity VR group. The low-intensity group received 40 min/day of VR for 3 days/week in addition to standard care. The high-intensity group received 60 min/day of VR for 5 days/week in addition to standard care. These results suggest that greater VR dosage could be associated with greater changes in function and movement.

3.5.1. Non-immersive Virtual Reality Technology

Many types of non-immersive virtual reality technology have been evaluated. Three studies^{9, 10, 13} utilized a computer game display and an arm orthosis. Figure 2 shows how a computer and arm orthosis are used together for VR treatment post-stroke.



Figure 2. *The VR and robotic TREX arm orthosis used by Housman et al¹³*

One study⁸ incorporated Nintendo Wii Sports into the VR group's treatment, while other researchers¹² employed Nintendo Wii gaming technology or a specialized rehabilitation gaming system (RGS) with a custom computer program and data gloves. Figure 3 depicts how a Nintendo Wii is used in clinical practice for stroke rehabilitation.



Figure 3. *Nintendo Wii Sports being used for stroke rehabilitation²²*

Reinforced feedback in a virtual environment (RFVE), which requires patients to grasp a real object with sensors while interacting with a screen display to complete tasks, was used by one study.¹⁴ Lastly, one trial²⁰ used a computer display with an attached joystick to enable subjects to move around in the digital

game. Since a variety of VR systems were tested in different studies, the effectiveness of VR must be interpreted with respect to the specific device used.

3.6. Immersive Virtual Reality Effectiveness

Authors have suggested that immersive VR technology may prove more beneficial than non-immersive VR because these provide “a combination of extra spatial transformation of uncoupled eye–hand movements, enhanced motion control, and more entertainment for patients, thereby motivating motion learning.”^{16, page}

¹ Fewer trials with high-quality evidence evaluate immersive VR compared to non-immersive VR because immersive VR is newer and more costly. Although one study¹¹ reported no between group statistical significance in motor function and strength, two studies^{17, 19} reported between group statistical significance favoring the VR group. One of these trials¹⁷ reported the VR group experienced significant improvement on the FIM test. The other study¹⁹ recorded significantly higher results on the FMUE, ARAT, FIM, and PASS in the VR group. As immersive VR technology develops, more testing will be required to determine if device sophistication impacts the effectiveness of VR for stroke rehabilitation.

3.6.1. Immersive Virtual Reality Technology

Various immersive VR technologies have been developed and examined. Two studies^{11, 19} utilized a VR headset, integrated games, and a device or program to track motion. Since these two studies^{11, 19} have similar VR technology, they demonstrate the impact of duration and intensity. Figure 4 shows a VR headset for stroke rehabilitation. One trial¹⁷ used PlayStation EyeToy games. This device projects a live feed of the user into a virtual environment on a wall-mounted display. This VR technology is defined as partially immersive. Since different types of immersive VR were used, the results and statistical findings must be interpreted within the context of the individual study and device used.

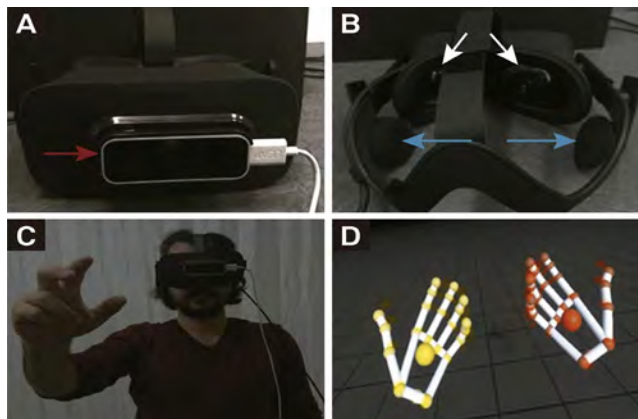


Figure 4. Virtual Reality Headset used by Ögün et al¹⁹

“General features of Leap Motion and virtual reality (VR) devices: A) Leap Motion (red arrow) mounted on a VR device; B) VR device; magnifying glasses (white arrows), headphones (black arrows); C) Hand Movements can be tracked without

using any devices on the hands via Leap Motion; D) All hand movements and the virtual environment also can be seen on a TV screen.^{19, page 3}

An ongoing study¹⁵ plans to enroll 60 participants, testing new, innovative VR technology: a VR headset, wireless hand-held controllers, movement tracking technology, and custom computer software. The trial¹⁵ will have a 3-week duration with high intensity (60 minutes/day x 5 days/week). The use of the newest and most advanced VR technology separates this study from the others; its equipment is more expensive and more sophisticated than that used in immersive VR trials to date. This study's¹⁵ results will be valuable to suggest whether the sophistication of VR technology impacts its effectiveness and to test the most modern applications with post-stroke patients.

4. Discussion

4.1. Appropriate Patients for Clinical Practice

Although existing studies suggest that VR can have a positive effect on patients with acute^{12, 14, 17}, subacute^{14, 17}, and chronic^{13, 14, 17, 19, 20} stroke, there are other factors that influence whether a patient is appropriate to receive VR. Since the equipment used for VR involves at least a software program and display, the cost can be relatively high. Non-immersive technology is less costly than immersive VR, but both forms require either the clinic or the patient directly to pay for equipment. Therefore, patients may be excluded if funding is lacking or if VR devices are not available to them. VR requires repetition and consistency to see results, whether done alone or combined with other treatments; so, it is best suited for motivated patients who are willing to commit themselves to VR sessions 3 times per week for at least 4 weeks. This treatment is not appropriate for patients who have visual impairments or eye disease because it requires use of a digital display. In addition, patients who easily experience nausea or vertigo with visual simulation should not participate in immersive VR. Further, patients who are pregnant or take medications that could cause headaches or dizziness should be further evaluated before treatment to ensure the VR will not cause adverse effects.

4.2. Comparison: Mirror Therapy

This section aims to compare VR intervention with Mirror therapy (MT) by examining the following factors: cost, inclusion and exclusion criteria, and effectiveness by construct. One meta-analysis, one scoping review, and three RCTs evaluating the effectiveness of Mirror therapy are included within this section.

4.2.1. Mirror Therapy Overview

Mirror therapy (MT) is a treatment that works by reflecting a patient's non-affected arm in a mirror simulating paretic arm movement.²⁴ Both MT and VR create a simulated reality for the patient to observe, and both originated in fields outside of stroke rehabilitation. VR was first invented for video gaming and

implemented for sports medicine; MT was originally created for phantom limb treatment. After years of research, both devices have become interventions used by patients post-stroke. For MT, a triangular prism-shaped mirror allows the patient to put the paretic arm inside of the mirror prism.²⁹ Patients are instructed to only look at the non-affected arm.²⁸ Figure 5 depicts the equipment used for MT. The movements executed by the non-affected arm include gripping fingers or external objects, flexing the elbow, moving arms from left to right, and lifting and lowering hands.²⁶ Similar to VR therapy, MT dosage is typically 20-60 min/day for 3 days/week and is often used with standard care.²⁵⁻²⁷ Three hypotheses suggest explanations for the underlying rationale of MT. One rationale explains that the engagement of the mirror neuron system (MNS) in the brain activates the corticospinal pathway (the main neural tract responsible for hand function). The second hypothesis suggests that learned non-use is targeted because motor pathways that are used by the non-affected arm can project to the paretic side of the body. Another rationale explains that increasing activities that activate visual parts of the brain can increase awareness of sensory feedback.²⁴



Figure 5. *The Triangular Prism-Shaped Device used for Mirror Therapy*³⁰

4.2.2. Comparison by Cost

While VR therapy requires expensive technological devices, MT only requires the mirror device and is much cheaper. VR equipment can range from a software program and display to a fully immersive VR headset and sensory gloves. Even though some VR treatments can be done at home,⁸ the majority of programs require the assistance of a therapist.^{9-14, 17, 19, 20} For MT, sensory objects such as beads or wooden blocks may be used for task-specific exercises, but the only required object is a mirror. Although the cost of MT is not insignificant, it is relatively inexpensive and can be performed from a clinic or at home.²⁸ To conclude, MT is less expensive and more convenient for patients to perform compared to VR intervention.

4.2.3. Comparison by Inclusion and Exclusion Criteria

Both VR and MT are highly accessible and have similar inclusion and exclusion criteria for appropriate patients. In section 4.1, the appropriate patients for VR intervention were examined and discussed. For MT, patients are not appropriate if they have severe upper arm pain, receptive aphasia, or spatial neglect.^{25, 26} Since

patients with spatial neglect ignore one side of their body, they may be unable to see the unaffected arm in the mirror's reflection.²⁵ Similar to VR, patients with visual impairments or eye disease may be inappropriate because patients need to view their unaffected arm in the mirror for the treatment to be effective.²⁵⁻²⁷ Both VR and MT have similar inclusion criteria and require that consenting patients have limited arm function²⁷ and intact cognitive function at baseline.²⁵⁻²⁷ The inclusion and exclusion criteria for MT subjects makes the treatment very accessible and is similar to that of VR therapy.

4.2.4. Comparison by Constructs

While some VR studies suggested it produced significantly better outcomes than standard care alone and was generally effective at improving arm motor function and grip strength,^{12-14, 17, 19, 20} other trials reported non-significant between group differences.^{8, 9, 11} No conclusion can be reached regarding the efficacy of VR intervention post-stroke.

Similar to VR, MT studies report mixed results. Although MT was shown to cause changes light touch perception and grip, the study²⁵ reported no significant differences in tests of motor ability and strength favoring the MT group. One scoping review,²⁴ with 28 trials included, suggests that MT causes significant improvements in fine motor control, dexterity, and tests of sensory loss. Another study²⁶ examined two experimental groups with MT and one control group; no statistically significant differences between unpaired MT and the dose-matched control group were reported. In contrast, one trial²⁷ reported significant improvement in spasticity and motor function when task-based mirror therapy (TBMI) was implemented. One meta-analysis²⁹ comparing MT and VR reported statistical non-significance in evaluating if MT was more effective than VR at treating individuals with arm dysfunction post-stroke; conversely, VR was not shown to be significantly more effective than MT. All in all, MT is similar to VR therapy in that no definite conclusion can be drawn regarding the efficacy of the treatment on certain outcome measures.

5. Conclusion

Virtual Reality (VR) may be effective in improving motor function, fine motor control, grip strength, and activities of daily living (ADLs).^{9-14, 17, 19, 20} Researchers have also demonstrated that VR can be effective in acute, subacute, and chronic stroke patients. A meta-analysis¹⁸ concluded that VR produced an average improvement of 28.5% of the maximal possible improvement in motor function and grip strength. VR with gaming produced significantly larger results (10.8%).¹⁸ Overall, more research is needed to determine optimal tools and dosage based on patient characteristics.

5.1. Unanswered Questions

5.1.1. What is the optimal intensity and duration for VR intervention?

In each of the trials evaluated in this review,^{8-14, 17, 19, 20} a unique duration and/or

intensity was used. Only one trial¹⁰ included VR groups using the same devices but at different intensities; however, no statistical significance was reported because of the small sample size. More studies should evaluate VR rehabilitation at different levels of intensity and compare the results using various outcome measures to determine how much VR is needed to see a significant amount of improvement compared to standard care. By only changing the intensity and duration, researchers would be able to rule out the effect of device type on the results. VR therapy is likely more effective at a higher intensity and duration because the patient will become more familiar with the device and make better use of the technology. The purpose of this experiment would be to find the critical point where the impact of VR therapy is the highest with the least amount of time needed.

Future studies should recruit patients and separate them into three dose-matched groups: 1) standard care only (control); 2) standard care and low-intensity VR; 3) standard care and high-intensity VR. Other variables, such as inclusion/exclusion criteria, device type, and stroke acuity, should be controlled between groups. In addition, standardized and comprehensive outcome measures should be used in future research trials to measure motor function, dexterity, arm movement, and fine motor control. These results would show how intensity and duration impact VR's effectiveness and what the optimal intensity is for stroke rehabilitation.

5.1.2. When is the best time to use VR post-stroke?

Although many studies^{8, 11, 14, 17} in this review included patients in more than one stage of recovery, none of the trials organized their groups based on stroke acuity. In addition, the research content and devices varied greatly, so no conclusion about the best time to use VR post-stroke can be reached.

VR therapy is likely to cause more physiological effects in acute and subacute post-stroke patients because the brain has more neuroplasticity and muscles have not atrophied, so they have more potential to regain function. Therefore, acute and subacute patients have a higher likelihood of seeing drastic functional gains than chronic stroke patients. However, chronic patients are more stable and often have more time to dedicate solely to repairing arm function. While acute and subacute patients often prioritize speech and leg function, chronic patients will be invested in regaining arm capacity. They likely will be more consistent with performing the VR arm therapy. Despite the difference in neuroplasticity, chronic patients may see a larger pre-post improvement on outcome measures compared to acute and subacute patients.

Further studies should recruit patients and separate them into two dose-matched groups. One group will receive experimental VR, and the other will receive sham treatment. Within each group, patient progress for each of the three acuities will be reported. Other variables, such as inclusion/exclusion criteria, duration, and intensity should be controlled between groups. In addition, standardized and comprehensive outcome measures should be used in future research trials to measure motor function, dexterity, arm movement, and fine motor control. Correlational statistics should be reported to determine if acute, subacute, or chronic patients saw the most benefit from the VR treatment. These results would be informative about when VR should be implemented into

rehabilitation to see the largest increase in outcome measures.

References

1. Stroke Facts. Centers for Disease Control and Prevention. <https://www.cdc.gov/stroke/facts.htm>. Published May 25, 2021. Accessed July 13, 2021.
2. Saposnik G, Levin M. Virtual Reality in Stroke Rehabilitation. *Stroke*. 2011;42(5):1380-1386. doi:10.1161/strokeaha.110.605451
3. Duncan Pamela W, Lai Sue Min, Keighley John. Defining post-stroke recovery: implications for design and interpretation of drug trials *Neuropharmacology*. 2000;39:835–841. doi:10.1016/s0028-3908(00)00003-4.
4. Katan Mira, Luft Andreas. Global burden of stroke in *Seminars in neurology*;38:208– 211Thieme Medical Publishers 2018. doi:10.1055/s-0038-1649503.
5. Ogourtsova AT. Virtual Reality – Upper Extremity. *Strokengine*. <https://strokengine.ca/en/interventions/virtual-reality-upper-extremity/>. Accessed July 14, 2021.
6. Laver KE et al. Virtual reality for stroke rehabilitation. *Cochrane Database Syst Rev*. 2017;11(11):CD008349. Published 2017 Nov 20. doi:10.1002/14651858.CD008349.pub4
7. Levin MF, Weiss PL, Keshner EA. Emergence of Virtual Reality as a Tool for Upper Limb Rehabilitation: Incorporation of Motor Control and Motor Learning Principles. *Physical Therapy*. 2015;95(3):415-425. doi:10.2522/ptj.20130579
8. Adie K, Schofield C, Berrow M, et al. Does the use of Nintendo Wii Sports™ improve arm function? Trial of Wii™ in Stroke: a randomized controlled trial and economics analysis. *Clin Rehabil*. 2017;31(2):173-185. doi:10.1177/0269215516637893
9. Byl NN, Abrams GM, Pitsch E, et al. Chronic stroke survivors achieve comparable outcomes following virtual task specific repetitive training guided by a wearable robotic orthosis (UL-EXO7) and actual task specific repetitive training guided by a physical therapist. *Journal of Hand Therapy*. 2013;26(4):343-352. doi:10.1016/j.jht.2013.06.001
10. Coupar FM. Exploring upper limb interventions after stroke. <http://theses.gla.ac.uk/3739/>. Published January 1, 2012. Accessed July 21, 2021.
11. Crosbie J, Lennon S, Mcgoldrick M, Mcneill M, Mcdonough S. Virtual reality in the rehabilitation of the arm after hemiplegic stroke: a randomized controlled pilot study. *Clinical Rehabilitation*. 2012;26(9): 798-806. doi:10.1177/0269215511434575
12. Cameirão MDS, Badia SBI, Duarte E, Verschure PF. Virtual reality based rehabilitation speeds up functional recovery of the upper extremities after stroke: A randomized controlled pilot study in the acute phase of stroke using the Rehabilitation Gaming System. *Restorative Neurology and Neuroscience*. 2011;29(5):287-298. doi:10.3233/rnn-2011-0599

13. Housman SJ, Scott KM, Reinkensmeyer DJ. A randomized controlled trial of gravity-supported, computer-enhanced arm exercise for individuals with severe hemiparesis. *Neurorehabil Neural Repair*. 2009;23(5):505-514. doi:10.1177/1545968308331148
14. Kiper P, Piron L, Turolla A, Stożek J, Tonin P. The effectiveness of reinforced feedback in a virtual environment in the first 12 months after stroke. *Neurol Neurochir Pol*. 2011;45(5):436-444. doi:10.1016/s0028-3843(14)60311-x
15. Huang Q, Wu W, Chen X, et al. Evaluating the effect and mechanism of upper limb motor function recovery induced by immersive virtual-reality-based rehabilitation for subacute stroke subjects: study protocol for a randomized controlled trial. *Trials*. 2019;20(1). doi:10.1186/s13063-019-3177-y
16. Saposnik G, Cohen LG, Mamdani M, et al. Efficacy and safety of non-immersive virtual reality exercising in stroke rehabilitation (EVREST): a randomised, multicentre, single-blind, controlled trial. *Lancet Neurol*. 2016;15(10):1019-1027. doi:10.1016/S1474-4422(16)30121-1
17. Yavuzer G, Senel A, Atay MB, Stam HJ. "PlayStation eyetoy games" improve upper extremity-related motor functioning in subacute stroke: a randomized controlled clinical trial. *Eur J Phys Rehabil Med*. 2008;44(3):237-244.
18. Karamians R, Proffitt R, Kline D, Gauthier LV. Effectiveness of Virtual Reality- and Gaming-Based Interventions for Upper Extremity Rehabilitation Poststroke: A Meta-analysis. *Archives of Physical Medicine and Rehabilitation*. 2020;101(5):885-896. doi:10.1016/j.apmr.2019.10.195
19. Ögün MN, Kurul R, Yaşar MF, Turkoglu SA, Avci Ş, Yıldız N. Effect of Leap Motion-based 3D Immersive Virtual Reality Usage on Upper Extremity Function in Ischemic Stroke Patients. *Arquivos de Neuro-Psiquiatria*. 2019;77(10):681-688. doi:10.1590/0004-282x20190129
20. Oh Y-B, Kim G-W, Han K-S, et al. Efficacy of Virtual Reality Combined With Real Instrument Training for Patients With Stroke: A Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation*. 2019;100(8):1400-1408. doi:10.1016/j.apmr.2019.03.013
21. Background Concepts in Stroke Rehabilitation. EBRSR. <http://www.ebrsr.com/evidence-review/3-background-concepts-stroke-rehabilitation>. Accessed July 26, 2021.
22. Wii gives the gift of restored movement and fitness to people after a stroke. Scimex. <https://www.scimex.org/newsfeed/wii-gives-the-gift-of-restored-movement-and-fitness-to-people-after-a-stroke>. Published April 11, 2018. Accessed July 25, 2021.
23. Assessments A-Z. Strokengine. <https://strokengine.ca/en/assessments/>. Accessed July 25, 2021.
24. Gandhi DB, Sterba A, Khatter H, et al. Mirror Therapy in Stroke Rehabilitation: Current Perspectives. *Ther Clin Risk Manag*. 2020; 16: 75-85.
25. Colomer C, Noé E, Llorens R. Mirror therapy in chronic stroke survivors with severely impaired upper limb function: a randomized controlled trial. *Eur J Phys Rehabil Med*. 2016; 52(3): 271-278
26. Choi HS, Shin WS, Bang DH. Mirror Therapy Using Gesture
27. Recognition for Upper Limb Function, Neck Discomfort, and Quality of Life After Chronic Stroke: A Single-Blind Randomized Controlled Trial. *Med Sci Monit*. 2019; 25: 3271-3278.

28. Madhoun HY, Tan B, Feng Y, et al. Task-based mirror therapy enhances the upper limb motor function in subacute stroke patients: a randomized control trial. *Eur J Phys Rehabil Med.* 2020; 56(3): 265-271.
29. McDermott AA. Mirror Therapy – Upper Extremity. *Strokengine.* <https://strokengine.ca/en/interventions/mirror-therapy-upperextremity/#Patient/FamilyInformation>. Accessed August 1, 2021.
30. Cabrales SX. Exploring the Effectiveness of Virtua Reality Therapy Compared to Mirror Therapy in Treating Individuals with Upper Extremity Dysfunction Following Stroke: A Meta Analysis. <https://scholarworks.calstate.edu/downloads/t435gf05k?locale=en>. Published May 2018. Accessed August 1, 2021.
31. Why Use Mirror Therapy? *Mirror Therapy.* <https://mirrortherapy.com/mirror-therapy/why-use-mirror-therapy/>. Published May 10, 2018. Accessed August 2, 2021.

Appendices

Appendix 1. Summary of Evidence

Author	VR Type	Time Since Stroke Onset	Summary of Treatment	Intensity and Duration	Outcome Measures	Findings
Adie et al. (2017)*	-At-home Nintendo Wii Sports (non-immersive)	-<=6 Months	-VR Group: Participants played Wii sports games in a seated position -Control: Participant-tailored arm exercises (based on the GRASP program) in a seated position	-Both groups warm up for 15 min and perform the intervention for up to 45 min/day for 6 weeks	-Assessed at baseline, 6 weeks, and 6 months -Action Research Arm Test -Canadian Occupational Performance Measure -Stroke Impact Scale -Modified Rankin Scale -EQ5D -Motor Activity Log Arm Function Test (6 months)	-No significant differences between groups on the Action Research Arm Test or Motor Activity Log after six months

Author	VR Type	Time Since Stroke Onset	Summary of Treatment	Intensity and Duration	Outcome Measures	Findings
Byl et al. (2013)⁹	-VR robotic orthosis (non-immersive)	->6 Months	<p>-Group 1: Performed bilateral tasks with device</p> <p>-Group 2: Performed unilateral tasks with device</p> <p>-Control: Repetitive exercises (reaching, grasping, self-care activities) with no dynamic arm orthoses</p>	-12 (90 min) sessions for a total of 18hrs of therapy	<p>- Assessed Pre/Post</p> <p>-Fugl Meyer</p> <p>-Motor Proficiency Speed</p> <p>-Box and Block</p> <p>-Tapper Test</p>	<p>- All patients (after 6 weeks) had significant improvements in flexibility, strength, pain, and voluntary movement (ULFM) ($p < 0.05$; effect sizes 0.49e3.53).</p> <p>-Each group significantly increased ULFM scores and range of motion</p> <p>-No significant between group differences</p> <p>-Virtual or actual Task Specific Repetitive Training (TSRT) with the robotic orthosis or physical therapists significantly reduced arm impairments (in the shoulder and elbow)</p> <p>-No significant increase in fine motor control in the hand</p>
Coupar (2012)¹⁰	-Armeo®Spring arm orthosis and VR games (non-immersive)	-<= 10 Days	<p>-Group 1: Conventional care</p> <p>-Group 2: Low intensity VR group</p> <p>-Group 3: High intensity VR group</p>	<p>-Low Intensity: Standard Care + 40min/day, 3days/week</p> <p>-High Intensity: Standard Care + 60min/day, 5days/week</p> <p>-Control: Standard care + Over Time for arm recovery</p>	<p>-Assessed Pre/Post and 3 months Post</p> <p>-Action Research Arm Test</p> <p>-Fugl Meyer UE</p>	<p>-There were higher changes in scores for the intervention groups (compared to the control group) between the baseline and the end of intervention assessments and baseline and the 3-month follow-up on the results relating to upper limb functional movement (ARAT), upper limb impairment (F-M), and level of disability (BL).</p> <p>-Statistical analysis was not recorded because of the limited number of participants, so conclusions cannot be drawn.</p>

Author	VR Type	Time Since Stroke Onset	Summary of Treatment	Intensity and Duration	Outcome Measures	Findings
Crosbie et al. (2011)¹¹	-VR mouse game, Headset, Gloves, and motion tracking (immersive)	-<= 24 months	-Group 1: Standard therapy group -Group 2: VR therapy group	-3-4-week period with a total of 9, 1-hour sessions	-Assessed Pre/Post and 6 months Post -The Upper Limb Motricity Index -Action Research Arm Test	-No significant between group difference -Both groups documented small changes to their upper extremity activity levels. -The Friedman test showed significant improvements within both groups as the number of sessions increased (the upper limb MI score (p = 0.008) and the ARAT score (p = 0.01)). -The Mann-Whitney U test showed no significant differences between groups. -7-8-point improvement for each group in the MI score -3-4-point improvement for the ARAT.

Author	VR Type	Time Since Stroke Onset	Summary of Treatment	Intensity and Duration	Outcome Measures	Findings
Da Silva Cameirão et al. (2011) ¹²	-Gaming system with data-collecting gloves and a visual display (non-immersive) -Wii (non-immersive)	-<= 3 weeks	-Group : Standard care + Specialized VR Rehabilitation Gaming System (RGS) -Group 2: Standard care + Wii gaming -Control: Standard care + Overtime working on motor skills (Dose Match)	-20 min, 3x per week, 12 weeks (12hrs).	-Assessed Pre, Week 5, Week 12, and Week 24 -Fugl Meyer -Chedoke Arm and Hand Activity Inventory	-The RGS group showed significantly improved performance in paretic arm speed and better performance in the arm subsection of the Fugl-Meyer Assessment Test and the Chedoke Arm and Hand Activity Inventory. -RGS group showed a significantly faster improvement rate for all the clinical scales during the treatment time frame.
Housman et al. (2009) ¹³	-Custom Software package with TREX arm orthosis (non-immersive)	->= 6 Months	-Group 1: experimental group -Control: conventional therapy (Dose Match)	-24 60 min sessions. 3x per week for 8 to 9 weeks.	-Assessed Pre/Post and 6 Months Post -Fugl Meyer -Rancho Functional test UE -Reaching ROM (deficit) -grip strength (dynamometer)	-There was a significant difference favoring the experimental group for the AMFM score at the 6-month follow-up -90% of experimental subjects reported a preference for VR therapy and recommended T-WREX over conventional therapy. -90% reported VR therapy "less boring, less confusing, and easier to track their progress than conventional tabletop exercises." ¹³

Author	VR Type	Time Since Stroke Onset	Summary of Treatment	Intensity and Duration	Outcome Measures	Findings
Kiper et al. (2011) ¹⁴	-Reinforced feedback in virtual environment (RFVE): subjects grasp a sensorized object while sitting in front of a screen to complete tasks (non-immersive)	-<= 1 year	-Group 1: 1 hour of standard care + 1 hour of RFVE -Control: Standard care for 2 hours (Dose Match)	-1 hour/day, 5x per week for 4 weeks (20 h)	-Assessed Pre/Post -Fugl Meyer -Functional Independence Measure - Modified Ashworth Scale (spasticity)	-A statistically significant between group comparison for the F-M UE and FIM (both $p < 0.001$), but not Ashworth ($p = 0.053$). -The RFVE therapy + standard care showed statistically better improvements compared to the standard care only. -The RFVE therapy + standard care was statistically more effective than the standard care double training for both post-ischemic and post-hemorrhagic groups.
Huang et al. (2019) ¹⁵ : <u>STILL UNDERWAY</u>	-VR system with a head-mounted display (HTC Vive-VR), a pair of wireless controllers, two base stations with steamVR® tracking technology, a computer, and a NVIDIA® Geforce® GTX 1070/MSI with 8 GB of GDDR5 (immersive)	->= 1 Week and <= 12 Weeks	-Group 1: 30 minutes of standard care + 30 minutes of VR therapy -Control: 60 minutes of conventional therapy (Dose Match)	-60 min/day, 5x per week, 3 weeks (15 h)	-Assessed Pre/Post, 12 Weeks Post -MRI Scanning -Fugl Meyer -Brunnstrom Stage Scores -Motor Assessment Scale	-Status: Subject Recruitment is taking place -The results of this study will provide evidence to either support or not support the continued investigation of immersive VR. - <i>This study is included because research on immersive VR to date is scarce. It is a study to watch continue tracking.</i>

Author	VR Type	Time Since Stroke Onset	Summary of Treatment	Intensity and Duration	Outcome Measures	Findings
Yavuzer et al. (2008) ¹⁷	-PlayStation EyeToy Games (Partly immersive because the game projects the user into the virtual environment, but there is no module or headset)	-<=12 Months	-Group 1: 60 min of standard care + 30 minutes of active VR intervention -Group 2: 60 min of standard care + 30 minutes of watching (nonactive) VR intervention without physically getting involved.	-30 min/day, 5x per week, 4 weeks (10 h)	-Assessed Pre/Post -Brunnstrom Stage Scores (UE and Hand) -Functional Independence Measure (FIM)	-The average score change (95% confidence interval) of the FIM reported significantly more improvement in the EyeToy group compared to the control group. -No significant difference was found between the groups for the Brunnstrom stages tests.
Öğün et al. (2019) ¹⁹	-VR Headset, Leap Motion Technology, and Integrated Games (immersive)	-> 6 Months and <= 24 Months	-Group 1: 60 minutes of VR therapy -Control: 45 minutes of conventional therapy + 15 minutes of passive/sham VR therapy (Dose Match)	-60 min/session, 3x per week, 6 weeks (18 h)	-Assessed Pre/Post -Action Research Arm Test (ARAT) -Functional Independence Measure (FIM) -Fugl-Meyer Upper Extremity Scale (FMUE) -Performance Assessment of Self-Care Skills (PASS)	-In the FMUE, ARAT, FIM and PASS tests, results were significantly higher in the VR group compared to the control -Both VR and conventional rehabilitation reported significant improvement in both upper extremity function and functional independence. A between group comparison showed that the VR group had significantly more improvement on tests for functionality, independence, and self-care.

Author	VR Type	Time Since Stroke Onset	Summary of Treatment	Intensity and Duration	Outcome Measures	Findings
Oh et al. (2019)²⁰	-VR combined with Joystim instrument (non-immersive)	-6 Months	-Group 1: 30 minutes of VR therapy -Control: 30 minutes of conventional therapy	-30min/day, 3x per week, 6 weeks	-Assessed Pre/Post and 4 weeks Post -Manual Muscle Test -Modified Ashworth scale -Fugl-Meyer -Hand grip -Box and Block -9-Hole Peg Test (9-HPT) -Korean Mini-Mental State Examination -Korean-Montreal Cognitive Assessment.	-Only the experimental group showed significant improvements in the MMT and MAS for wrist extension, MAS for elbow flexion, and Box and Block Test scores. -Only the experimental group showed significant improvements in the 9-HPT score, the lateral, palmar, and tip pinch power scores in the post-training period when compared to the pre-training period. -The experimental group had significantly improved levels of spasticity of elbow flexion 4 weeks post training.

Appendix 2. Outcome Measures

Outcome Measure Test	Assessment Type
Action Research Arm Test (ARAT)	Measures Limb function/Ability to handle different objects/Activity limitation
Canadian Occupational Performance Measure	Patient reported assessment of daily function, self-care, and productivity
Stroke Impact Scale (SIS)	Patient reported assessment of strength, memory and thinking, hand function, communication, Activities of Daily Living/Instrumental Activities of Daily Living (ADL/IADL), emotion, mobility, and participation
Modified Rankin Scale (MRS)	Measures degree of disability and level of dependence in daily living
EQ5D	Measures quality of life (dependence, mobility, discomfort, depression)
Motor Activity Log (MAL)	Measures upper extremity performance in real life situations
Fugl Meyer Upper Extremity Test (FMUE or FM or FMA)	Measures motor function, sensory loss, balance, joint pain, and joint range of motion
Motor Proficiency Speed Test	Measures fine motor control, strength, and coordination
Box and Block Test	Measures gross unilateral manual dexterity
Tapper Test	Measures sensory loss, coordination, and fine motor control
The Upper Limb Motricity Index	Measures arm and hand strength
Chedoke Arm and Hand Activity Inventory	Measures motor recovery/control and grip strength
Rancho Functional Test- Upper Extremity	Measures motor function and skill level performing functional tasks
Reaching Range of Motion (ROM) Test	Measures movement, coordination, and flexibility
Dynamometer Test	Measures hand grip strength
Functional Independence Measure (FIM)	Measures independence in performing daily tasks and motor control
Modified Ashworth Scale	Measures muscle spasticity

Outcome Measure Test	Assessment Type
Brunnstrom Stage Scores	Measures motor function and spasticity
Motor Assessment Scale	Measures motor function, movement, arm strength, and arm tightness
Performance Assessment of Self-Care Skills (PASS)	Measures capacity for Activities of Daily Living/Instrumental Activities of Daily Living (ADL/IADL)
Manual Muscle Test	Measures muscle strength and arm function
9-Hole Peg Test (9-HPT)	Measures finger dexterity and fine motor control
Korean Mini-Mental State Examination	Measures cognitive abilities
Korean-Montreal Cognitive Assessment	Measures cognitive abilities



Opening Our Hearts: An Analysis of How Noel Pearson Reframes the Debate through Irony and Identification

Laura Zhang

Author Background: *Laura Zhang grew up in Australia and attended Ravenswood School for Girls in Sydney, NSW in Australia. Her Pioneer research concentration was in the field of sociology/political science/history and was titled “Rhetoric and Social Movements.”*

1. Introduction

At the core of Australian political rhetoric is the ever-evolving Indigenous Australian social movement. Indigenous Australians rallied in a civil rights and land rights movement from the mid-1960s to the late-1990s to appeal for self-determination policies, those that ensure “the right of all peoples to freely determine their political status” (Australians Together) and the creation of a system for Indigenous Australians to reclaim their land. Recently, as shown by the conception of the 2017 Uluru Statement from the Heart, a convention requesting the “establishment of a ‘First Nations Voice’ in the Australian Constitution” (AIATSIS), this movement has shifted to call for the constitutional recognition of Indigenous communities (Nakata, 335-354). Despite decades of advocacy, systemic injustice and inequality have prevented the attainment of such recognition, making this movement as relevant now as ever.

An orator who provides critical insight into the history of the Indigenous Australian movement is Noel Pearson, an Indigenous lawyer and lands rights activist who also founded the Cape York Institute for Policy and Leadership (Indigenous Voice). Although there has been substantial analysis on non-Indigenous parliamentary speeches such as former Prime Ministers Paul Keating’s “The Redfern Address” (1992) and Kevin Rudd’s Apology Speech (2008), not only is there a disappointing lack of analysis surrounding Indigenous Australian rhetoric but there has been a lack of analysis surrounding Pearson’s speeches, a gap this paper hopes to fill. Moreover, as Pearson’s speeches come from a bottom-up and outside-the-system background, they are integral to understanding the Indigenous perspective on the issues this movement aims to address.

This paper will focus on Pearson’s 1996 speech, “An Australian History for Us All” which was delivered at the University of Western Sydney. Through this speech, Pearson engages with two significant contemporary events: the abolishment of *terra nullius* and the recent rhetoric of former Prime Minister John Howard. By analyzing this speech, this paper hopes to demonstrate how

Pearson uses the master trope of irony and the rhetorical technique of identification to criticize Howard's rhetorical use of "hot button issues" such as "white guilt" and reframe the political discourse surrounding the Indigenous Australian social movement.

To explore how Pearson achieves identification with his audience through irony to reframe the debate, this paper will apply Kenneth Burke's Theory of Identification. It will demonstrate how Pearson overcomes his constraints of persuasion by creating identification through a humble opening, then how he establishes the "consubstantiality" between the Australian public and the Indigenous movement through dialogical irony, and finally how he defines the rhetorical situation. It will ultimately argue that through reframing the debate, Pearson mobilizes the audience to fight for the democratic demands of the Indigenous Australian social movement by "opening our hearts a little" (Pearson).

2. Literature Review

Aboriginal and Torres Strait Islander Indigenous Australians have lived in Australia for over 65,000 years (National Museum Australia) and have a deep, spiritual connection to their land. In 1778, when the British colonized Australia, they superseded Aboriginal sovereignty under the justification of *terra nullius*—that Australia was "nobody's land"—hence they had the legitimacy to take ownership of Australia. *Terra nullius* was rooted in the presumption that Aboriginal peoples were "savages... like madmen" (Hunter, 61), a paternalistic model that constructed Indigenous Australians as the "other" and colonization as a catalyst for modernization. Consequently, Aboriginal peoples have been subject to a long history of discriminatory assimilation policies, exemplified by the Stolen Generation, a term describing the forcible removal of Indigenous children from their families to integrate into broader Australian society under the pretense that it would improve their quality of living in the 1910–1970s (Australians Together). Moreover, when the Commonwealth of Australia was established in 1901, Indigenous Australians, under the Australian constitution, were excluded from voting, pensions, enlistment into the war, and land rights, privileges to which the rest of Australia were entitled (SBS).

These policies catalyzed the creation of the Indigenous Australian social movement, centered around a democratic demand—individual demands addressed within an institution—for land rights and self-determination (Laclau, ch.4). Through advocacy, the High Court of Australia made the "historic decision" that Eddie Mabo, a Torres Strait Islander, owned legal titles to the land he and his family had presided on. Also known as the *Mabo v Queensland* trial, this ruling established the *Native Title Act 1993* and abolished the legal doctrine of *terra nullius* (Burgmann, 23-26). After being adopted by the Australian government, this effectively established a law that allowed for an Indigenous person to possess ownership of their land, fulfilling the demands of the movement. Resultingly, there was a transition in the historical discourse surrounding Indigenous Australians from "The Great Australian Silence" to the incorporation of the Indigenous perspective and the realities of colonization (Beasley, 6-7). "The Great Australian Silence" is a term coined by anthropologist W.E.H. Stanner in *The*

Dreaming & Other Essays to describe the pre-1970s “cult of forgetfulness” which resulted in the omission of the Indigenous experience in Australian historical discourse, subjugating Indigenous Australians into a “melancholic footnote” (11).

Despite progress, a countercultural, conservative backlash spearheaded by John Howard emerged, criticizing this “black armband” view of history, demonstrating a “tension” (Ernesto, ch.4) between the demand for self-determination and the institutional order. The “black armband,” a term coined by historian Geoffrey Blainey, describes the abandonment of the “optimistic view of history” by focusing solely on the negatives of Australia’s past such as the past treatment of Indigenous Australians (McKenna) instead of the “good character of the nation” (Beasley, 5). This highlights how there was a newfound polarization between Indigenous Australians calling for recognition and self-identity, and institutions that relied upon mythologies designed to create an edifice of white guilt during the 1970s-1990s. The notion of “white guilt,” which stems from a fear of the acknowledgement of historical injustices (Steele, 501-502), is often used by Pearson to antagonize dominant institutions and demand self-determination—a request that has not been previously met. This is the point in the Indigenous Australian social movement where Pearson performed his speech.

Although there is an abundance of historical literature surrounding the Indigenous Australian social movement, there has been a lack of rhetorical literature analyzing the discourse of Indigenous activists. In the existing literature, comparisons have often been made between Pearson and Stan Grant (1986-present), an Indigenous journalist and writer. In 2015, Grant presented a viral speech as part of the IQ2 debate series at the Ethics Centre on the topic “Racism is Destroying the Australian Dream,” noting that “[t]he Australian Dream is rooted in racism. It is the very foundation of the dream” (Grant). Scholars such as Helen Razer highlight the similarities between Pearson’s attempt to reframe the debate and Grant’s calls for reconciliation, referring to Grant’s rhetoric as “[a]n elaboration of Pearson’s” (Razer). Grant himself acknowledged in an article for *The Guardian Australia* that “brilliant orators and thinkers like Noel Pearson... have inspired me beyond words” (Grant).

Despite similarities, Sarah Payne notes that “Grant steps away from the guilt associated with the ‘black armband view of history’” (Payne) which Pearson readily engages with, whilst journalist Amy McQuire attributes the success of Grant’s speech to his “non-threatening diplomacy” (McQuire). Thus, Pearson is identified as a foundational orator, whose rhetoric serves as an influence for other Indigenous activists, making him a representative figure in the Indigenous Australian social movement. Simultaneously, this paper disagrees with the argument that Pearson “tell[s] what white Australia... wants to hear” (McQuire). Rather, it argues that Pearson’s rhetorical uniqueness derives from his willingness to engage in discourse about controversial topics like “white guilt” whilst striving to not alienate or admonish his audience, instead aiming for identification.

Regarding Pearson’s speeches, there has not been a formal analysis of the rhetoric in “An Australian History for Us All,” which makes this paper the first to do so. This essay will attempt to fill this gap in theoretical discourse, as well as a holistic perspective of Pearson’s rhetoric and more broadly, the demands of the Indigenous Australian social movement.

There has been substantial theoretical literature around Burke’s Theory

of Identification and the master trope of irony, the rhetorical techniques this paper will use to analyze Pearson's speech. As articulated by Diane Davis, Kenneth Burke was heavily influenced by Sigmund Freud's theory of identification and this "anxiety of influence" led him to create a post-Freudian definition of identification, or "consubstantiality," as a mode of "symbolic action" that transcends the natural division between audience and orator by unifying the two parties on a commonality. (Davis, 123-147) In *A Rhetoric of Motives* (1969), Burke furthers the traditional Aristotelian interpretation of rhetoric by claiming that persuasion can only occur after stylistic identification, when an orator establishes rapport and a sense of common identity between him and the audience, lending significance to the technique of identification. (46) Moreover, central to his definition of identification is its ironic counterpart: division. (Burke, 23) Burke argues that identification arises and is earnestly "affirmed" due to division (Jackson, 7-8). As Burke wrote *A Rhetoric of Motives* shortly after World War II, he equally emphasized the importance of retaining one's individuality and "locus of motives" whilst being consubstantial with another, as there is a danger in total identification—the "ultimate disease of cooperation: war" (Burke, 22).

His discussion of the interplay between division and identification has resulted in literature exploring the mutual influence between irony and identification. Kenneth Burke refers to irony as one of the four master tropes, along with metaphor, "a device for seeing something in terms of something else," metonymy, and synecdoche (Burke, *Four Master Tropes*, 421). Irony, as expressed by Quintilian, is "a figure of speech or trope in which something contrary to what is said is to be understood" (qtd. Vlastos, 21). This can be further distinguished between verbal and situational irony, the types of irony that Pearson applies most frequently throughout his speech.

Verbal irony, as defined by Connon Thirlwall in "On the Irony of Sophocles," is "a figure which enables a speaker to convey his meaning with greater force by means of a contrast between his thought and expression" (qtd. Hutchens, 355). For example, a speaker declaring "what nice weather!" in the middle of a storm intends to underscore his dislike for the weather through a contradictory statement, creating verbal irony. David Wolfsdorf further clarifies how verbal irony occurs when a speaker is transparent with his true intentions, succeeding when the audience understands grasps "the speaker is highlighting the literal falsity of the utterance" (175). A further type of verbal irony elucidated by the likes of Gregory Vlastos and Iakovos Vasiliou is Socratic irony. In *The Dialogues of Plato*, Socrates is represented as an *eirōn*, who disavows himself of knowledge, feigns ignorance, and humbly questions the *alazōn*, or his interlocutor (Vlastos, 86). This mask of innocence exposes the fallacies in the interlocutor's arguments, leading to the truth—resulting in irony in the contrast between the *eirōn*'s appearance and the reality of the discourse.

In comparison, situational irony is a contrast where a "situation either leads to its opposite or is endowed with some opposite significance" (Nathan, 40). For instance, a thirsty man stranded in the middle of the ocean can be viewed as situational irony from the incongruence between the setting, a body of water, and the man's inability to quench his thirst, or what he is expected to do. In rhetorical terms, situational irony is predicated on an audience's ability to juxtapose the orator's statements with the rhetorical situation, understanding how his statements contrast with the speaking situations.

Beyond a rhetorical trope, irony can function as a dialogue, “resulting when two or more perspectives enter dialogue with one another” (Murray, 29). For instance, when a speaker ironizes and questions another way of seeing through linguistic strategies, he may dissociate the audience from this way of thinking, creating a dialectic between audience and speaker in search of new meanings. In such cases, the orator has increased identification with the audience and reframed the debate. (Parson). In David Kaufer’s *Irony and Rhetorical Strategy* (91-105), he demonstrates the effect of irony on identification as irony can:

1. Create affiliation by identifying a common victim.
2. Victimize an audience while addressing it.
3. Keep incompatible audiences at bay.
4. Represent a false character (*eiron*).
5. Be epideictic in pursuit of group cohesion.

However, Kaufer also articulates how, for irony to be effective, the speaker must make an audience understand his true intentions and overcome the persuasive constraints of the rhetorical situation, or else the meaning cannot be communicated (91-105). Thus, there is an interplay between irony and identification: irony may cause identification, but for irony to be used appropriately, the speaker must establish consubstantiality between him and the audience first.

In an attempt to fill in the gap in literary discourse on the rhetoric of the Aboriginal Australian social movement and the progression of Noel Pearson’s rhetoric, this paper examines how Pearson utilizes the interplay between identification and irony to reframe the political discourse surrounding the Indigenous Australian social movement and mobilize the audience to fight for Indigenous self-determination.

3. Analysis

3.1 Defining the Debate

Pearson’s use of identification and irony in “An Australian History for Us All” will be analyzed in terms of how he challenges the institutional backlash against Indigenous self-determination and reframes the debate. On November 20th, 1996, Noel Pearson gave the speech “An Australian History for Us All” to an academic gathering at the University of Western Sydney. As an outside-the-system figure, his views and perspectives are distinct from his audiences, creating a division. To overcome this persuasive constraint, he starts the speech with a humble opening, saying how he comes “only with some observation about how our popular understanding of the colonial past is central to the moral and political turbulence we are still grappling with as Australians” (Pearson). The formal register indicates his awareness of his distinguished audience, intentionally establishing himself as the humble *eiron*, allowing for the potential use of Socratic irony. The collective pronouns “we” and “our” compounded with his reference to the commonality between him and the audience—that both are “Australians”—create consubstantiality. The understatement of “some observation” creates a modest tone, allowing Pearson to elevate the role of the audience above his own to

increase his ethos. Thus, Pearson effectively establishes a rapport between him and the audience.

Once he has established his role as the *ieron* and identified the commonalities between him and his audience, Pearson subtly incorporates Socratic irony to suggest his true intentions to the audience. Immediately after his humble opening, he says that he “fears” he is conducting “what Prime Minister John Howard calls the perpetual seminar for elite opinion about our national identity” but “will nevertheless persevere” (Pearson). The juxtaposition between Howard’s description of “elite opinion” to his previous humble tone creates a clash between Pearson’s established perspective that his speech will be a simple social commentary and Howard’s perspective that any discussion about Australia’s past is “elite.” The ironic tone of “persevere” mocks Howard’s dramatization of the discourse around Australian identity and establishes Pearson’s intentions of dissociating his audience with Howard’s rhetoric. In painting Howard as the *alazon*, Pearson sets up a Socratic dialogue to uncover the flaws in the political status quo. Here, Pearson’s outside-the-system and bottom-up position in the rhetorical situation allows him to openly criticize Howard’s perspective—something that would have political ramifications if he were inside the system, allowing for effective criticism.

He continues to dissociate the audience from the status quo by defining the current political discourse. This is demonstrated when he articulates that the “guilt about Australia’s colonial history... is a hot button issue in the Australian community” (Pearson). The metaphor comparing the portrayal of “guilt” to a “hot button,” an American colloquialism referring to issues that elicit strong emotional responses, imbues the term with properties of melodrama, allowing Pearson to implicitly criticize Howard’s antagonistic portrayal of guilt in the media.

He situates the audience in the status quo by saying “[y]ou would not need to be a political genius to bet that the guilt issue is one of the keenest buttons that the Federal Member for Oxley and her followers in our national government have pressed and with great electoral resonance” (Pearson). The allusion to Pauline Hanson, the “Federal Member for Oxley” who was, at the time, a notorious advocate for the abolition of government assistance for Indigenous Australians, further imbues the issue of white guilt with implications of racism as an intentional impediment to affirmative action. The litotes of “not need to be a political genius” implies through ironic understatement that the antonymous phrase, “need to be a political genius,” is untrue, therefore highlighting the contrary—that any layperson can understand that this guilt issue is a “[keen] button.” Through verbal irony, Pearson highlights how the guilt issue should be apparent to the audience as nothing more than a “hot button.” His ambiguous tone expresses his indirect criticism of the Australian public resonating with this issue despite it being, in his perspective, a construct of an anti-Aboriginal Australian agenda.

Pearson further states that “John Howard implies that history has been cultivated by the politically-correct classes which urges guilt and shame upon Australians about the national past” (Pearson). Pearson criticizes the disparaging usage of “political correctness,” which conventionally refers to the caution taken against using words that might be insulting to people from oppressed minorities (Hamilton), to demonstrate how Howard has mischaracterized progressive inclusivity as a mechanism for guilt and shame. Instead, he challenges the view

that progressive politics has cultivated guilt about the past and offers the view that it has created, in Howard's words, "a responsibility to put an end to the suffering" of Indigenous people (qtd. Pearson). By creating an ironic Socratic dialogue between Howard's rhetoric and his own, Pearson defines Howard's—and other contentious political figures—as disparaging and antagonistic, dissociating the audience from the status quo and framing the debate on white guilt as intentionally harmful to the Indigenous Australian social movement.

3.2 Dissociating the Audience

After defining the political status quo, he clarifies his stance within the context of the demands of the Indigenous Australian social movement, further dissociating the audience from Howard's rhetoric. He states that the "Anglo-Celtic story of Australia's past was seriously distorted by significant omissions and by some straight out fictions, such as the fiction of 'peaceful settlement.' The certitude with which history and the humanities generally proclaimed the myth of *terra nullius* meant that the legal invisibility of Aboriginal people and a steadfast belief in our inhumanity was embedded into popular belief" (Pearson). The verbally expressed quotation marks around "peaceful settlement" mock the mainstream historical narrative that reflects positively on European colonization, demonstrating Pearson's perspective that settlement was, ironically, nothing but violent. The cumulative repetition of "fiction" and "myth" compounds Pearson's stance that doctrines such as *terra nullius* which aim to absolve Australia's violent history by creating "belief in our humanity" are fabrications that obscure the truth of European subjugation. The visual imagery evoked by "legal invisibility" alludes to the concept of "The Great Australian Silence" and the omission of the rights of the Indigenous within the Australian political discourse. Thus, Pearson clarifies his perspective that there should be no ambiguity to the injustices Indigenous Australians faced in Australia's history and identifies a counter-issue to white guilt: the preclusion of the Indigenous voice in Australian discourse.

After identifying this issue, he calls out the hypocrisy in the actions of greater Australia to redirect the debate away from the issue of white guilt through the rhetorical question:

For how can we as a contemporary community in 1996 share and celebrate in the achievements of the past, indeed feel responsibility for and express pride in aspects of our past, and not feel responsibility for and express shame in relation to other aspects of the past? (Pearson).

The audience is shown the situational irony of the Australian veneration of Australian history and the hypocritical omission of the "shame" of colonial dispossession within the Australian historical narrative, which serves as a claim against institutional neglect of responsibility for the past. He then reminds the audience that "[a]fter all, the heroic deeds at Gallipoli and Kokoda are said to be ours as well. Lest we forget" (Pearson). The historical allusion to "Gallipoli," a military operation conducted during World War I involving 16,000 Australian and New Zealand soldiers, which is regarded as "the foundation legend of Australian military history" (National Museum Australia) uses the collective consciousness of his audience to highlight the absence of the Indigenous

perspective within the Australian history, creating consubstantiality. The slogan “lest we forget,” which has become symbolic of the commemoration of the sacrifice of Australian soldiers in conflicts, (Anzac Portal) ironically juxtaposes Australia’s intentional remembrance of their military past to their hypocritical forgetfulness of the colonial oppression of Indigenous peoples. Therefore, through situational irony, Pearson’s admonishes “black armband history” and Howard’s criticism of the movement’s supposed “allocation of guilt” as such rhetoric may lead to the harmful institutional preclusion of the Indigenous voice in Australian political discourse.

3.3 Reframing the Debate

Pearson then starts to reframe the debate by creating identification with the audience. He does so by drawing attention to the division between the “ordinary Australians” who are not “still afflicted with the obscurantist tendencies of the past... and are prepared to acknowledge the legacies of the past” and the so-called guilty Australian (Pearson). Through the evocative tone of “afflicted,” Pearson compares white guilt to a disease, motivating his audience to dissociate with the term and identify idealistically with the hypothetical group of “ordinary Australians,” for which Pearson implicitly advocates for. To create further identification and reframe the debate towards the necessity of relinquishing guilt on an institutional and societal level, Pearson characterizes the “other.” This is demonstrated when he states that “[t]he more vehement the denials” of “some responsibility to deal with the legacy of the past in the present,” “the more they betray and anxiety to exorcise guilt” (Pearson). The verbal irony in the word “exorcise” comparing guilt to a demon demonstrates the over-extent to which “they”—those afflicted with white guilt—are so defensive of “their own identity and heritage” (Pearson, *White guilt*) that they are compelled to evade responsibility and action to ameliorate the current status of Indigenous peoples. Instead, he implies that the “optimum position for non-Indigenous people” (Pearson, *White guilt*) is an acknowledgement of the pain of dispossession, abandoning white guilt and working to amend this reality, therefore reframing the debate.

After creating identification, Pearson reminds the audience of how relinquishment of this debate led to actual change, advocating for the acknowledgement of self-determination. This is encapsulated in his statement, “Mabo threw the country into social, political and psychological turmoil. I always said that it was the turmoil and confusion the country had to have” (Pearson). The intertextual allusion to then-treasurer Paul Keating’s statement about the Australian recession given during a speech on November 29, 1990, as a “recession Australia had to have” ironically frames the abolishment of *terra nullius* due to the Mabo case as a revolutionary decision that had the same social, political, and legal ramifications as those of the recession (Keating). The verbally ironic interplay between “turmoil and confusion” and Pearson’s view that it “had to have” happened subtly criticizes the hyperbolized Australian reaction to the Mabo case as, in Pearson’s view, it should have been the expected outcome. As Keating was also a proponent of Indigenous self-determination, Pearson then implicitly establishes the link between the necessity for the acquisition of land-rights to the need for self-determination, establishing ethos by using the rhetoric of another

orator and creating identification within the audience through allusion.

Consequently, Pearson demonstrates the irony of the rhetorical situation to rationalize his call for unity and acknowledgement of Indigenous dispossession and reframe the debate. This manifests towards the end of his speech where he says that the media debate of the “hot button” slogans of “[b]lack armbands. Guilt Industry. Political Correctness” has created “brain-damaged dialogue” (Pearson). The deprecating metaphor “brain-damaged” continues to dissociate the audience from the current political discourse. The cumulative listing of the slogans and metaphor of “hot button” derides Howard’s criticisms of “black armband history,” devaluing the slogans’ importance in the justification of the denial of reconciliation. Here, Pearson demonstrates how he has stripped these phrases of their loaded connotations and properties in the audience’s eye, and how he has reframed the debate away from these slogans.

He furthers this dialogue with the hypophora: “[i]n conclusion: what substance is there in the new emphasis on our colonial history that Prime Minister John Howard and his Minister are urging and in the crusade against the back armbands and their alleged obsession with guilt? The answer is nothing at all” (Pearson). The hyperbolic undertones of a religious “crusade” mirrors the use of the word “exorcise” earlier in the speech, associating the hypervigilance of Howard’s rhetoric against the slogan of “white guilt” to a religious charge to highlight the absurdity of the rhetorical situation. The posing of a question and then answering it straight away mirrors that of a conversation, where the audience is the questioner, allowing Pearson to assume the role of the trusted, undisputed authority on this matter. Thus, the audience is convinced of the situational irony that despite the actions of the media and the government to combat the “black armbands and their alleged obsession with guilt,” there is no “substance,” or necessity, in doing so.

3.4 Opening our Hearts

Resultingly, Pearson dissociates the entire Indigenous Australian social movement from guilt to reframe the debate towards the need for self-determination. Instead, Pearson offers the view that “it is not about guilt, it is about opening our hearts a little bit” (Pearson), again alluding to Keating’s rhetoric when he called upon Australia to “open our hearts” in *The Redfern Address*. The colloquialism “opening our hearts” (Keating) evokes an emotive tone, contrasting his previous ironic rhetoric to offer the straightforward sentiment that the Indigenous social movement does not demand division and retribution, but reconciliation and unification. In doing so, he creates a dialectic between him and the audience, urging them to identify with his cause and realize the exigency of the situation. Pearson criticizes public political discourse to highlight the situational irony of the divide between the Australian public, institutions, and the Indigenous social movement, showing how reconciliation is “not about guilt” (Pearson). By highlighting how reconciliation should be a mutual aim for all parties, Pearson creates identification in the audience, urging them to advocate similarly for Indigenous self-determination by “opening our hearts.”

He continues the symbol of the open heart by saying “...an open heart means that if a people have suffered wrongs and have wounds that are still keen,

then there must be some respect for that” (Pearson). The imperative “must” creates a hortative tone, didactically guiding his audience to the right course of action—demonstrating respect, rather than guilt, for Indigenous Australians. The visual imagery of “wounds” evokes pathos, targeting the audience’s emotions and demonstrating the present-day ramifications of colonization on Indigenous peoples. This ultimately mobilizes the audience to strive for the acknowledgement of colonial oppression and the inclusion of the Indigenous voice in Australian historical discourse.

Finally, Pearson concludes his speech through intertextual allusion to reframe the debate one last time. Pearson states that he is “sure that Robert Hughes, whose seminal work *Culture of Complaint* was touted as the first foray against political correctness... would be ashamed to see what is passing for free speech and history in this country today” (Pearson). The appeal to authority of Robert Hughes invites the audience to attach Hughes’ conception of “guilt” as a “potent justification for the plunder, murder and enslavement of people” (qtd. Pearson) to their understanding of the political discourse, reframing the debate in the minds of the audience. Thus, through using irony and the technique of identification, Pearson successfully characterizes the political discourse surrounding the Indigenous Aboriginal movement, dissociates the audience from Howard’s slogans of political correctness, black armband history and white guilt, and reframes the debate to advocate for the acknowledge of colonial history, mobilizing the audience to “open our hearts” (Pearson).

4. Conclusion

This paper has analyzed Noel Pearson’s speech “An Australian History for Us All” through the lens of Kenneth Burke’s Theory of Identification and the master trope of irony. First, Pearson uses a humble opening and establishes his ethos to create identification with his audience and signify his true intentions to use verbal and situational irony to analyze the debate. The use of Socratic irony in “An Australian History for Us All” criticizes Howard’s use of the hot button catchphrases of “black armband,” “guilt” and “political correctness,” dissociating the debate from the question of the allocation of guilt to that of how self-determination and equality can be achieved for Indigenous Australians. In doing so, he characterizes the fight for constitutional recognition as something that demands the attention of all Australian citizens urgently, challenging the status quo and calling for the public to “open our hearts” and mobilize against the institution. Thus, he addresses the situational irony in the division between the Australian public and the Indigenous Australian social movement, calling for unification and identification based on shared ideals and through “opening our hearts” (Pearson). In the context of the Indigenous Australian social movement, as Pearson’s speech was one of the first to address Howard’s rhetoric and delivered at an esteemed academic gathering, it can be argued that it took use of the rhetorical situation and catalyzed the shift to reframe the debate.

As this paper is one of few treatises that have focused specifically on the Indigenous perspective and voice in regards to the Indigenous Australian social movement, it has both social and scholarly significance. This paper serves as the

first analysis of Pearson's "An Australian History for Us All" and can serve as a foundation for further analysis of the speech or otherwise a series of his works. Moreover, this paper has used the concept of irony and identification in tandem—further scholarly research can utilize the framework of their mutual influence to analyze the effect of irony and identification on an orator's rhetoric.

There are three main limitations to this paper. First, it characterizes Pearson as a representative of the Indigenous Australian social movement. Despite his influence on orators such as Stan Grant, he is often viewed as a divisive figure in Aboriginal communities. As an article from *Sydney Morning Herald* states, Pearson has been scrutinized as one of "those most strongly criticizing the new order" but "failing their community" (Waterford). More broadly, as there are over 500 different groups of Indigenous Australians in Australia, (Australian government) it is difficult to generalize the aims of Noel Pearson as representative of the Indigenous community as a whole. Further research on a range of Indigenous speakers and their rhetoric can overcome this limitation. Second, by limiting the scope of this paper to irony and identification alone, this paper cannot cover the breadth of rhetorical devices used by Pearson in his speeches. Future scholars can aim at providing a more holistic view of Pearson's rhetoric. Third, by focusing on a singular speech, this paper is unable to provide a holistic overview of Pearson's rhetoric. Future scholars should attempt to analyze a series of his speeches to chart the evolution of Pearson's rhetoric and the Indigenous Australian Social Movement as a whole.

Works Cited

- Australian, Government. "Our people." Australia.gov.au, n.d., <https://info.australia.gov.au/about-australia/our-country/our-people>. Accessed 04 July 2021.
- Australians Together. "The Indigenous civil rights movement in Australia." Australians Together, n.d., <https://australianstogether.org.au/discover/australian-history/civil-rights-movement/>. Accessed 12 June 2021.
- Australians Together. "The Stolen Generations." Australians Together, n.d., <https://australianstogether.org.au/discover/australian-history/stolen-generations>. Accessed 14 June 2021.
- Beasley, C. "The Breaking of the 'Great Australian Silence': How and Why the Writing of Indigenous Australian History Has Changed over the Last 40 Years." *The ANU Undergraduate Research Journal*, 5. 2014. <http://press-files.anu.edu.au/downloads/press/n1673/pdf/Caroline-Beasley.pdf>. Accessed 20 June 2021.
- Burgmann, V. "The Aboriginal movement in power, profit, and protest: Australian social movements and globalization". Kooriweb. 2013. <http://www.kooriweb.org/foley/resources/pdfs/13.pdf>. Accessed 19 June 2021.
- Burke, K. *A Rhetoric of Motives* (Rev. ed.), Berkley and Los Angeles, California, University of California Press, 1969, communication cache, http://www.communicationcache.com/uploads/1/0/8/8/10887248/keneth_burke_-_a_rhetoric_of_motives_1950.pdf. Accessed 20 July 2021.

- Burke, K. "Four Master Tropes." *The Kenyon Review*, vol. 3, no. 4, 1941, pp. 421-438. JSTOR. <https://www.jstor.org/stable/4332286>. Accessed 14 June 2021.
- Davis, D. "Identification: Burke and Freud on Who You Are." *Rhetoric Society Quarterly*, vol. 38, no. 2, 2008, pp. 123-147. Jstor, <https://www.jstor.org/stable/pdf/40232541.pdf?refreqid=excelsior%3a068c89a053b10f7abc136628a6b59c8d>. Accessed 22 July 2021.
- DVA (Department of Veterans' Affairs) "'Lest we forget' as a symbol of commemoration in Australia". DVA Anzac Portal, n.d. <http://anzacportal.dva.gov.au/commemoration/symbols-commemoration/lest-we-forget-symbol-commemoration-australia>. Accessed 23 June 2021.
- Grant, S. "Australia Day: What makes us great, and what that greatness demands of us." *The Guardian*, 25 January 2016, *the guardian*, <https://www.theguardian.com/commentisfree/2016/jan/25/the-overwhelming-response-to-my-speech-tells-me-we-can-be-better-than-our-racist-history>. Accessed 13 January 2022.
- Grant, Stan. "Racism and the Australian Dream". Address at the IQ2 debate at the Ethics Centre. 2015. Sydney. Speech.
- Hamilton, C. "Political Correctness: Its Origins and the Backlash Against it". *The Conversation*, 2019, <https://theconversation.com/political-correctness-its-origins-and-the-backlash-against-it-46862>. Accessed 06 July 2021.
- Hutchens, E. N. "The Identification of Irony." *ELH*, vol. 27, no. 4, 1960, pp. 352-363. JSTOR, <https://www.jstor.org/stable/2872064>. Accessed 11 January 2022.
- Indigenous Voice. "Mr. Noel Pearson." Australian government, n.d., <https://voice.niaa.gov.au/whos-involved/biography/mr-noel-pearson>. Accessed 13 June 2021.
- Jackson, Chonticia Y., "The Use of Rhetoric in Public Relations: Kenneth Burke's Theory of Identification and Consubstantiality" (2013). Dissertations, Theses and Capstone Projects. Paper 556.
- Kaufer, David. "Irony and Rhetorical Strategy". *Philosophy & Rhetoric*, vol. 10, no. 2, 1977, pp. 90–110. JSTOR. www.jstor.org/stable/40237017. Accessed 21 June 2021.
- Keating, Paul. "The Redfern Address". Launch for the International Year of the World's Indigenous People. 10 December 1992. Redfern Park. Speech.
- Laclau, E. *The Rhetorical Foundations of Society*. Edited by E. Laclau, Verso, 2014, Perlego.
- McKenna, M. "Different Perspectives on Black Armband History". 5. Politics and Public Administration Group. Parliament of Australia, 1997, https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/RP9798/98RP05. Accessed 19 June 2021.
- McQuire, A. *The Viral Rise of Stan Grant: Why Diplomacy Won't be Enough for Our People*. 2016, <https://amymcquire.wordpress.com/2016/03/25/the-viral-rise-of-stan-grant-why-diplomacy-wont-be-enough-for-our-people/>. Accessed 13 January 2022.

- Mills, S. "I am sorry': Prime ministerial apology as transformational leadership". *Studies in Australian Political Rhetoric*, 2014, https://press-files.anu.edu.au/downloads/press/p291051/html/ch01.xhtml?referer=&page=8#toc_marker-9. Accessed 20 June 2021.
- Murray J W. "Kenneth Burke: A Dialogue of Motives." *Philosophy & Rhetoric*, vol. 35, no. 1, 2002, pp. 22-49. Jstor, <https://www.jstor.org/stable/pdf/40238108.pdf?refreqid=excelsior%3ad174cd44bbadbb582ba45d6f46c9095c>. Accessed 20 July 2021.
- Nakata, S. "Who is the Self in Indigenous Self-determination?" *Indigenous Self-Determination in Australia*, 2020, pp. 335-354. Doi: 10.2307/j.ctv1bvncz1.21. Accessed 13 June 2021.
- Nathan, A. R. *Plato's Use of Irony*. Diss. Thesis, The University of Sydney, 2020, pp. 35-42.
- National Museum Australia. Evidence of first peoples. Nma.gov, Australian Government, 2021, <https://www.nma.gov.au/defining-moments/resources/evidence-of-first-peoples>. Accessed 14 June 2021.
- National Museum Australia. Gallipoli landing. National Museum Australia, 2021, <https://www.nma.gov.au/defining-moments/resources/gallipoli-landing>. Accessed 27 June 2021.
- O'Sullivan, D. (2021, February 26). "Indigenous recognition is more than a Voice to Government - it's a matter of political equality". *The Conversation*. <https://theconversation.com/indigenous-recognition-is-more-than-a-voice-to-government-its-a-matter-of-political-equality-154057>. Accessed 20 June 2021.
- Parson, D. W. "ISSA Proceedings 2006 – Rhetorical Irony: Problems and Possibilities." *Rozenberg Quarterly Magazine*, 2006, <https://rozenbergquarterly.com/issa-proceedings-2006-rhetorical-irony-problems-and-possibilities/>. Accessed 22 July 2021.
- Payne, S. "Stan Grant's 'Racism is Destroying the Australian Dream' speech." *NEW: Emerging Scholars in Australian Indigenous Studies*, vols. 2-3, no. 1, 2018, pp. 84-86. *UTS ePRESS*, <http://dx.doi.org/10.5130/nesais.v2i1.1481>. Accessed 13 January 2022.
- Pearson, N. "White guilt, victimhood and the quest for a radical centre." *Griffith Review*, <https://www.griffithreview.com/articles/white-guilt-victimhood-and-the-quest-for-a-radical-centre/>. June 2007. Accessed 30 June 2021.
- Pearson, Noel. "An Australian History for Us All". Address to the Chancellor's Club Dinner. 20 November 1996. University of Western Sydney. Speech.
- Razer, H. "Helen Razer on Stan Grant's 'Australian Dream'." *Review of Racism and the Australian Dream*, by S. Grant, *Daily Review*, 3 December 201, <https://dailyreview.com.au/helen-razer-stan-grants-australian-dream/>. Accessed 13 January 2022.
- Rose, S. "How the Word 'Woke' Was Weaponised by the Right". *The Guardian*, 2020, <https://www.theguardian.com/society/shortcuts/2020/jan/21/how-the-word-woke-was-weaponised-by-the-right>. Accessed 06 July 2021.
- Soutphommasane, T. "A Brave Piece of Legislation!: The Racial Discrimination Act, 40 Years On. Australian Broadcasting System", 2015. <https://www.abc.net.au/religion/a-brave-piece-of-legislation-the-racial-discrimination-act-40-year/10098574>. Accessed 27 June 2021.

- Stanner, W. *The Dreaming & Other Essays*. Black Inc. Agenda, 2010. <https://ereader.perlego.com/1/book/969325/0>. Accessed 20 June 2021.
- Steele, S. "White Guilt." *The American Scholar*, vol. 5, no. 4, 1990, pp. 501-502.
- Sydney Broadcasting Service. "Timeline: Indigenous Rights Movement." SBS News, 2021, <https://www.sbs.com.au/news/timeline-indigenous-rights-movement>. Accessed 20 June 2021.
- The Australian Institute of Aboriginal and Torres Strait Islander Studies. "Eddie Koiki Mabo." AIATSIS, Australian Government, n.d., <https://aiatsis.gov.au/explore/eddie-koiki-mabo>. Accessed 12 June 2021.
- Vasiliou, I. "Conditional Irony in the Socratic Dialogues." *The Classical Quarterly*, vol. 49, no. 2, 1999, pp. 456-472. *JSTOR*, <https://www.jstor.org/stable/639871>. Accessed 12 January 2022.
- Vlastos, G. "Socratic irony." *The Classical Quarterly*, vol. 37, no. 1, 198, pp. 79-9. *JSTOR*, <https://www.jstor.org/stable/639346>. Accessed 12 January 2022.
- Waterford, J. "Noel Pearson is a great orator but he's essentially a leader without followers in Aboriginal world." *The Sydney Morning Herald*, 07 November 2014, <https://www.smh.com.au/opinion/noel-pearson-is-a-great-orator-but-hes-essentially-a-leader-without-followers-in-aboriginal-world-20141107-11fq9f.html>. Accessed 5 July 2021.
- Wolfsdorf, D. "The Irony of Socrates." *The Journal of Aesthetics and Art Criticism*, vol. 65, no. 2, 200, pp. 175-178. *JSTOR*, <https://www.jstor.org/stable/4622222>. Accessed 12 January 2022.
- University of Sydney Library Sydney. (2003). *An Historical Journal of the Transactions at Port Jackson and Norfolk Island* [PDF]. University of Sydney Library. <https://adc.library.usyd.edu.au/data-2/hunhist.pdf>. Accessed 14 June 2021.



Trans Selfhood, (Mis)recognition, and Post-Identity Resistance

Moxuan Xiong

Author Background: *Moxuan Xiong grew up in China and currently attends Rutgers Preparatory School in Somerset, New Jersey in the United States. Their Pioneer research concentration was in the field of philosophy and titled "The Examined Life."*

Abstract

How do we understand the trans subject in relation to medical histories that have created conditions for affirmative self-transformations but also normative discourses that seek to pathologize gender-variant bodies? This paper aims to examine the conditions and limitations of trans narratives within the established confines of identity politics. It questions the political and epistemic stakes implicated in such formulations of trans selfhood and how forms of representation initiated by the proliferation and enforcement of normalizing identity categories tend to enact a genre of medical discourse that has been historically formative in reifying and erasing trans subjectivities. A trans historiography that does justice to its subjects, I believe, should be critically descriptive of the struggles faced by trans individuals and in turn suggests the precondition for alterity and freedom. Thereby, I theorize beyond accounts of trans selfhood that center around ontological discordance as a form of strategic essentialism and instead outline a vision of trans resistance as currently entangled in a web of intersecting struggles. Departing from this vision, I argue for a framework of recognition as a question not of identity but social status, so that racial, ethnic, sexual, gendered, classed, and other kinds of differences can be engaged in a convergent, rather than divisive, dynamics of shared statuses, shifting the focus from identity-based self-realization to a picture of self-creation that aligns with the needs and desires of trans people without recourse to restrictive categories that function by the hierarchizing grammar of intelligibility.

Introduction

Called by an injurious name, I come into social being, and because I have a certain inevitable attachment to my existence, because a certain narcissism takes hold of any term that confers existence, I am led to embrace the terms that injure me because they constitute me socially. The self-colonizing trajectory of certain forms of identity politics are symptomatic of this paradoxical embrace of the injurious term.¹

Enacted as an identity category, “trans” is fraught with discursive baggage from the very beginning, starting with the pathological theorizing around the trans subject and moving toward the mobilization of trans identities as a site of radical gendered possibilities. Disembodied accounts of trans selfhood, while centering trans subjects as the locus of discourse, foreclose the very possibility of trans subjectivity by invalidating trans experience either as dislocated deviance that demands corrective assimilation or as a contingent vehicle for political polemics. But, first of all, what does the trans experience entail, setting aside what is given as the singular truth? How do trans becoming and resistance unfold as emerging from contextual forces that seek to police and regulate certain identities as recognizable and others not?

To begin with, I outline the conditions of trans medical discourse from which predominantly recognized accounts of trans selfhood are enacted. Then, I argue that the identity-based reduction of trans selfhood to an account of self-realization and representation not only erases the complexity of individuals’ desires, experience and self-understanding, it is inadequate for a vision of trans future as it entrenches normalizing forms of identity that ultimately sustain, rather than challenge, inequality, oppression and domination. Finally, I explore the intersections of gendered struggles to suggest the possibility of moving forward toward multiple visions of trans futurity and away from the traps of internecine, identitarian conflict.

The Medical Discourse of Transsexuality and Its Discontents

Narratives of trans selfhood in mainstream contexts tend to act on a model of self-realization grounded in an innate gender identity, which takes the idea of a true self as inherently stable and, for its internal facticity, consistently regulates attempts at self-transformation toward an authentic gendered self (Mason-Schrock, 1996). The “wrong-body” model, paradigmatically deployed as a validating account of trans selfhood, frames the misalignment between gender identity and the sexed body within a medically mediated account of the gendered self—an invisible, innate identity which is to be realized through a series of medical procedures that operate within certain prioritized definitions of sex/gender throughout trans medical history².

¹ Judith Butler, *The Psychic Life of Power*, 104.

² This involves, for instance, the development of sex reassignment surgery (SRS) which centers treatments around genital-centric models of sex and the newer procedures of facial femi-

Such a narrative, developed from the discursive practices of the medical establishment which sought to pathologize trans identities and gender variance—indeed, instituted these particular forms of gendered existence as pathological, continue to shape modes of self-narrative in established relations between the trans subject (patient/customer) and the medical community (representing the normalized means of recognition through which the diagnosed subjects become women/men, hence their authentic selves)³. As such, the gendered self is innate and stable, wherein authenticity is inwardly realized by fixing its misalignment with a “wrong body.”

Medical discourses informed by (and largely enacting) such an account then constitutes a genre within which certain trans identities are formed and represented, often in pathologizing terms that further configure the institutionalization of the intertwined relationship between diagnosis and demand, which structure existing norms regarding trans subjects and hence accounts of trans selfhood, making some recognizable while excluding others that do not live up to those viably accounted modalities as institutional and discursive practices shape forms of belief, identities and self-knowledge (Stone, 1991; May, 2002).

It should be noted that, while the medical establishment is closely related to and, to some extent, historically constituent of the emergence of trans identities, the medical discourse should not rule out any possibility of perceiving trans resistance. Attempts that foreclose such a possibility, exemplified by Raymond⁴ and Hausman, are falsely entrapped in a totalizing, reductionist picture that assimilates all possible modes of trans personhood into to a hetero-patriarchal medical discourse of transsexuality, whereas what calls for critical examination is precisely the discursive structure through which one-dimensional reductionist pictures are automatically ascribed to trans subjects.

Double Vision, Double Dilemma

While established medical discourse that implicates the “wrong-body” model imposes a pathological imperative, the question lies: in what terms should the medical account be received in relation to trans identities and particularly trans agency? In other words, should the medical account be taken as a definitive prescription that denies any theorizing beyond the medicalized and, therefore, reified non-subject, or, should it be read as critically descriptive of the existing struggles faced by trans individuals, suggesting a precondition for a different account that opens the doorway for the coming-to-voice of trans people?

nization surgery (FFS) which is focused on the face as a gendered site of intersubjective recognition (Plemons, 2017).

³ The patient-doctor relationship, while collaborative, is also one of conflicts, given the prevailing tendency on the part of health-care providers, psychiatrists especially, to pathologize the expression of trans feelings, leading to adversarial relationships between doctors and patients and reinforcing an already paternalistic attitude toward trans people’s subjectivities (Plemons & Straayer, 2018).

⁴ Janice G. Raymond did particular injustice with her *Transsexual Empire*, published in 1979, which drew upon anti-transsexual discourses that had been developing since the 70s and argued that trans people were the mindless agents trapped in the false consciousness of patriarchal oppression bent on the denigration of women (Stryker, 2004).

My answer clings to the latter reading, that is, that the medical discourse is a condition of subjectivation, not a preclusion of it. Attempts to theorize transsexuality within the medical framework only work to further pathologize the conditions from which trans subjects necessarily emerge, creating an impasse for any alternative formulations.

In *Changing Sex: Transsexualism, Technology, and the Idea of Gender* (1995), for example, Bernice Hausman sets out to offer a feminist critique of transsexuality within a Foucauldian framework that addresses the medical discourse of transsexuality. Curiously, while deconstructing the medical regime through which trans subjects are constituted and hence pathologized, Hausman reaffirms this picture by foregrounding the entire dependency of transsexual agents upon the medico-technological system. Within this picture, trans subjects are defined by their very desire for medical interventions, which in turn constitute them as such—a necessarily pathological mode of existence. This way, trans autobiographies only serve to justify access to medical resources, as trans subjectivity is formed entirely through medical accounts of transsexuality. Beyond the medical model, no subjectivity is possible at all. While Hausman might rightly identify a picture of trans dilemma enmeshed in pathologizing discourse, it sees the dilemma as the single reality which rules out all possibilities of trans subjectivation, a picture that falsely overlooks the potentiality of resistance and alterity within existing conditions of transness.

Indeed, alterity has always been present, as forcefully voiced in Sandy Stone's *The Empire Strikes Back: A Posttranssexual Manifesto* (1991), which sought the closing of predominant medical and feminist accounts of transsexuality to disclose a multiplicity of trans narratives. This reflexive gesture toward the medical model thus provides a point of departure for theorizing variant visions of possibilities, making the precondition for a formulation of trans selfhood in its troubled reality.

Returning to “wrong-body” model, not as a recourse to some reifying medical gaze but simply to reexamine the pervasive logic of such an account closely, we may nevertheless identify a dissonance between “felt body” and the external appearance of the body that is often characteristic of trans experience. Far from being pathological, the dissonance between the felt sense of body and external appearance of the body should be seen as a response to the working of existing gendered systems that produce intelligible and unintelligible feelings, experiences, and attitudes. As Gayle Salamon notes, “the production of normative gender itself relies on a disjunction between the 'felt sense' of the body and the body's corporeal contours and this disjunction need not be viewed as a pathological structure” (Salamon, 2010). While this felt dissonance might be ascribed to an internal, ahistorical logic of gendered identification, it powerfully reflects the historical and normative forces that render it distinctly urgent and palpable. This experience of “wrong” embodiment then engages in the plurality of processes of forming the trans subject, wherein complications immediately arise: it is in experiencing dissonance and the ensuing feelings, attitudes, and beliefs, that the subject is rendered ambiguous and illegible. The very feeling of dislocation that frequently characterizes such an experience should illuminate the processes through which the deviant subject is called out, addressed in terms that defy apparent legibility.

(Mis)recognizably Trans

*One will be a person of integrity—whole and of a piece, someone to be counted on, stable and steady; one's beliefs, attitudes, and feelings will be explicable and coherent; one's actions will follow straightforwardly from one's intentions; one will be simultaneously solid and transparent—a block of unclouded substance.*⁵

While the narrative of trans subjectivation within an inwardly regulatory framework might be ascribed to an individualist, self-enclosed account of identity in a critical gesture, the social presumption that such an ascription enacts nevertheless risks reducing trans selfhood to a given embodied collectivity as part of a homogenizing project while dismissing the constructed, intersected dimension of the gendered self. From this presumption would then arise a transgender paradigm that endorses a politics of recognition installed to assimilate trans bodies into a narrative of inclusivity, so that trans subjects, through an accommodative apparatus of recognition, become intersubjectively visible and viable as normalized bodies.

Dominant multicultural models of recognition posited by theorists such as Taylor and Honneth approach the core of injustice as deriving from a distorted form of recognition, a “recognition deficit” that is to be remedied by improving existing, problematic strategies of recognition while leaving its mechanics untouched (Taylor, 1994; Honneth, 1995). Models of recognition as such, which ground identity-based self-realization upon normative assumptions for reciprocal recognition that underlies intersubjective relations, enact an ethics of authentic individual development that, rather than resolves the core of trans misrecognition as entangled in a mesh of interlocking issues, further subscribes to the discursive project that produces viable and unviable identities.

Underlying identity models of recognition is an insoluble tension between two visions: on one hand, it is to secure for the marginalized and oppressed the access to benefits and privileges that are necessary for being recognized as “fully human”, an intelligible, respectable member of society; on the other, it is to disrupt privileged models of selfhood and ideals of normalcy, that is, what it counts to be fully human, so that one is promised the right to recognition and the choice of non-conforming (Scheman, 2011). Yet the promise is false, as the choice to deviate is but an empty rhetoric when deviation is cast as a prerequisite for entering the gilded arena of recognition. In recognizing the value of distinctive traditions and modes of life, the institutional authority necessarily centers certain modalities as dominant so as to differentiate the deviant others, who, for their unrecognizability, are prompted to gather around group-identities that manage to be different in a regulated (or what is called “authentic”) way in order to remain viable. This way, ameliorative recognition grants the now legible individuals rights and benefits that define a good, viable life, worthy of love, esteem, and respect. The double enforcement that seeks to “normalize difference” is self-defeating. Securing certain archetypes of difference as the validating cornerstone of a political identity inevitably privileges some modes of subjectivity and representation (one needs to perform an authentic, intelligible selfhood in order to be representable) by excluding others. The agency of cherishing

⁵ Naomi Scheman, *Queering the Center by Centering the Queer. Shifting Ground*, 111.

difference, as Charles Taylor argues, is underlay by the dialogical character of human life: "We become full human agents, capable of understanding ourselves, and hence of defining our identity, through our acquisition of rich human languages of expression. The genesis of the human mind is in this sense not monological, not something each person accomplishes on his or her own, but dialogical." (32). Here, it is evident that an ontological threshold is being posited for an identity politics of recognition: before cherishing one's distinctness, it must be granted that the person be a "full human agent," capable of exhibiting those differences in a triumphant gesture and therefore speaking a common language that is richly assured and intelligible. It does not take into account the scene of address before the dialogue, where the speaking agents are situated in imbalanced relations to one another, and the one who is really speaking must also be the one who looks down at and tries to understand the unintelligible other who is undertaking to perform dialogical agency so as to be recognized. Recognition secured as such is informed precisely by misrecognition: the one who is addressed to give an account of themselves, fails to do so, for the price of that account is a renunciation of the very agency of becoming⁶—it is the variable opacity and unintelligibility that mark the constitution of the self (Butler, 2005).

It is also through this opacity that the self remains in an ever-shifting web of relations to one another, where intersections of struggle and oppression take hold⁷. The imposition of authentic self-knowledge is therefore impossible and only does violence to the unintelligible subaltern when models of solidly transparent subjectivity are centered around what is already viably recognized. Hence, as recognition works along a singular, drastically simplified group-identity in the matrix of intelligibility, one is prompted to offer an account on behalf of who one is, or, more crucially, a collective Self that elides multiplicity and alterity for the enactment of its own legibility. The gesture of self-narrating as such thus locates the struggle in representation—the internalized dislocation is to be replaced by new, authentic self-representations collectively enacted to produce a distinct culture, whereby respect, and esteem are promised (Fraser, 2000).

Identity models of recognition that seek to legitimize an authentic collective of self-representations turn out only to be counterproductive, for the cost of being recognized as deserving of personhood, rights, and citizenship demands adherence to those very impossible standards. It requires that one be different enough to complement the self-congratulatory spectacle of multicultural difference, but only in ways that solidify existing conceptions about the kinds of subjects qualified for the benefits of political membership. What the self-generative model of trans selfhood fails to address, then, is the conflicted dimension of social and political becoming that institutes "trans" as a collective identity, in which process institutionalized practices that seek to define and configure the subject are enforced. In this light, the social relations in which embodied accounts of internal displacement are enacted operate by the same logic that subjects deviant bodies to violence and repression.

⁶ When a diagnosis is required for access to trans medicine, rather than "I want to become a woman/man," a trans person is expected to justify themselves by saying, "I was always a woman/man inside and I need my body to match." This scene of address also precludes the possibility of describing oneself as belonging to non-binary sex-gender categories as gender-queer people may also seek medical transformations.

⁷ This is a point that underlies my argument for a post-identity politics of intersectional coalition, which will be addressed in a later section.

Tying the trans subject to a collective of internally authentic and self-generative identities, the identity model of recognition dismisses—and further discourages—the multiplicity and complexity of the processes through which trans selves are formed and lived in their differing specificity⁸.

The Suffering Subject and The Case of Trans Surgeries

A powerful narrative deployed to legitimize trans experience and identity for access to medical treatment and legal justifications is one of suffering. As the World Professional Association for Transgender Health (WPATH) puts it in the Medical Necessity Statement:

The medical procedures attendant to gender affirming/confirming surgeries are not “cosmetic” or “elective” or “for the mere convenience of the patient.” These reconstructive procedures are not optional in any meaningful sense but are understood to be medically necessary for the treatment of the diagnosed condition. In some cases, such surgery is the only effective treatment for the condition, and for some people genital surgery is essential and life-saving (Knudson et al. 2016: 3).

Here, trans surgeries are rendered medically necessary for what is termed gender dysphoria (GD) according to Diagnostic and Statistical Manual of Mental Disorders (DSM), which posits an internal, etiological incongruence between the sexed body and the gendered self that should be reconciled through medical interventions. While the need for gender-affirming health care is common among trans people, for which a diagnosis of GD is often imperative, it paradoxically serves to assist trans affirmative care while constraining trans subjects’ self-narratives. This constraint is embedded in the constitution of trans selfhood with the medical discourse of suffering—one needs to be an adequate sufferer of GD to qualify for trans medicine and other interventions, which, in this context, are “essential and life-saving.” This narrative must also require that the subject, through necessary medical interventions, be brought back to an inherent, invisible identity, by which gender is formulated as essential, fixed, predefined, and beyond the control of the subject (Prosser, 1998). The discourse of suffering which frames trans selfhood with the imperative of suffering also assumes a compulsory gender binary that marks ambiguity as taboo and undesirable. Further, the appeal to suffering is strategically exploited when it comes to legal rights to medical opportunities for transgender youth: since trans children are already one of the most vulnerable groups with high rates of suicide and depression, life-saving trans medicine should be secured in order to alleviate mental problems while they experience GD. The differential rhetoric of exclusive vulnerability

⁸ For an ethnographic inquiry into the identitarian category of transgender and how it produces misrecognition, see David Valentine’s *Imagining Transgender*, Chapter 3, where he discusses the application of trans categories to individual lives and how the use or non-use of the category entangles the terms in which they have become institutionalized, producing these selves as unintelligible. This in turn shows how the aims and logics of identity politics themselves produce this evident unintelligibility and elide an analysis of the established inequalities that underlie them when, for instance, white, upper middle class versus non-white, lower class trans people find it easier to align their self-understandings with discrete categories of identity while others are hard-pressed in doing so (Valentine, 2007).

(wherein GD is seen as a distinctly trans experience and an essential marker of trans identities) initiates a purchase of political psychology intertwined with a normative telos of one-dimensional identity: only a specific style of suffering will count, as the subject's capacity to construe and act upon its body is limited by the medical gatekeeper (Heyes & Latham, 2018). The morbidity of this strategic schematization is revealed when strategy becomes enforcement. As suffering comes to define trans narratives, it also individualizes gender and disallows critique of the systems that contribute to trans people's suffering in the first place. It erases the variant complexities of trans experience, including gender euphoria, relief with an ambiguous anatomy, embrace of an inconsistently gendered life, or critique of the psychopathological systems that discipline trans subjects.

Another example is the pervasive imperative of "passing" or, more exclusively, credentials of gender confirmation surgeries⁹, which are widely necessary in mainstream and most legal contexts for someone to become recognized as legitimately trans (women/men). Trans identities that do not seamlessly fit in the bifurcated gender system, as a result, are further excluded from the ruling domain of recognition. Specific forms of recognition, as shifting in accord with the trans medical establishment and the operations of sex/gender systems that institute them in the given picture, may be improved toward normalized models of embodiment that smoothly assimilate trans bodies into existing modes of identification. Facial feminization surgery (FFS), for instance, is now deployed often as an improved, complementary strategy of intersubjective recognition versus the genital-centered sex reassignment surgery (SRS) for effecting more successful experiences of embodiment, enabling trans women to become better recognized as their true selves and working along an account of authentic self-realization (Plemons, 2017). However, the current mechanics of power by which such recognition works¹⁰ necessarily produces legitimate trans identities (white, unambiguously recognizable as women/men, economically privileged to an extent that makes FFS and others measures of ameliorative recognition affordable) by reproducing existing sex/gender norms and excluding those that fail or refuse to assimilate, making them illegible and susceptible to varied kinds of violence. The "truly embodied self" that trans subjects are directed toward in this project of authentic self-realization are nevertheless repressive¹¹, not because the individual pursuits are endorsing a politically unfavorable

⁹ Not just surgeries per se—in scenes of address one must disclose the "truth" about one's sex, whereas surgical outcomes are seen as deceptive and only the document will be truthful about the person in question. It is also contended that one does not simply "pass" because to say that someone passes as a woman/man is to confirm that woman/man is not a category to which she or he truly belongs, creating a double standard in which the gender work of cisgender people is buttressed as natural and undemanding whereas trans people's gender is retained as factitious and willful.

¹⁰ Again, not the desirability of the surgery itself (the legitimacy of creative self-fashioning should be granted) but the machinery through which it is offered as inherently desirable, which serves as the same mechanics of power that produces subjects of failure—by forging a certain archetype of authenticity as intrinsically necessary, it enforces an impossible standard that continually reifies and ostracizes the failing and deviant.

¹¹ It should be noted that availability of trans medicine remains a crucial goal for trans struggles. My contention is to make room for the possibility that securing equitable treatment must not require trans people to undertake pathologizing (self-)interpellation by appealing to exclusive, reifying discourse of suffering, and that safeguarding diagnosis is not the only gateway to acquiring trans medical opportunities.

project, but because the way those pursuits are forged and positioned implicate them in a scene of injustice where structures of violence, oppression and subordination are enforced into the fabric of trans becoming.

As advocates and governing bodies of trans medicine make trans surgeries including FFS a medical necessity dissimilar to cosmetic surgeries (the demand for which, argued as such, does not suffice the psychological seriousness that makes trans surgeries a medical necessity), the assumption in turn confirms the unviability of trans women who cannot “pass” because of unwillingness or inability to go through FFS when in fact both are enacting a normative understanding of the sex-gender regime (Heyes & Latham, 2018). On the other hand, one can get a sense of how cosmetic surgery is stigmatized in this account as ethically undesirable in relation to trans surgeries because it forms a scene of deception—however “feminine” they look, women who undergo cosmetic surgeries can never be authentic—that is, “real women”—because they are acting upon a schematism of deception with recourse to medical intervention. By making FFS a necessity, this logic of deception is not done away but is rather inextricably installed into the notion of trans selfhood *per se*—whether one does it or not, one is deceiving (Bettcher, 2007). The regulation of viable trans identities and representations hence presents the trans self as an ontological failure in the first place—unless it “detransitions” back to normalized categories, a narrative that conservative and anti-trans media frequently deploy against trans people (Knox, 2019). As the current political condition for trans subjectivation is predominantly distinguished by difficulty or unwillingness to identify with and seamlessly observe a sex/gender regime which dictates the possible forms of recognizable identities available, by setting out to expand the relations of recognition to those not receiving sufficient respect and esteem, the identity model is reinforcing the very failures that are to be addressed.

A Note on Inclusivity

While the relationship between the political and medical apparatus of gender transition and trans people is seen as reciprocal (which might also be maliciously phrased as “complicit” in an anti-trans gesture to stigmatize trans subjectivities), it should also be remembered that such a relation is framed in concrete encounters that are historically shaped in response to an insoluble hostility. The neoliberal account of “becoming who you are regardless of what they think of you” is impossible. It is also true—acutely so—when one considers the medical procedures that are deemed intrinsically desirable for trans people seeking transition, the end of which should yield a form of recognition that is viable and likable. There are moralistic contentions within queer political circles that denounce procedures that serve the purpose of “passing” including FFS as erasing trans visibility and foreclosing the possibility of a trans specific radical political subjectivity or assert that trans men and women should go beyond the gender binary to remain subversive (Bornstein, 1994; Bettcher, 2014). Not coincidentally, the same moral rhetoric may arise from trans-exclusionary feminist voices that pose growing threats to trans communities because recourse to medical transformation, in their account, necessarily reinforces the gender stereotypes that perpetuate oppression of women, in which definition trans women are invariably excluded. What is not touched upon in this accusatory diction is the fact that individual action is contextually restricted in the first place. Transmisogyny assails the logic of desirable forms of self-realization which are structurally imposed upon trans individuals, including those who find it difficult

conforming to or simply deny the demanding, strictly bifurcated gender norms, the very system which seeks to render trans identities unintelligible and unviable at the outset of the story. It is not the case that we are already in a gender-free society where some antisocial rascals decide to “be themselves” by disrupting the binary order of cisheterosexual patriarchy (the existence of which already problematizes the gendered categories we live by) like a carefree rebel (though rebellion is always gestured toward our liberation). This picture is historically empty and does brutal injustice to subaltern trans people in how they experience their lives—endless encounters with threats, disgust, violence, and misrecognition. The assimilationist agenda, on the other hand, creates an impossible pattern by proposing a series of gendered and social endeavors in order for trans people to be recognized (be wealthier, more beautiful and hard-working, etc). Neoliberal inclusivity procedures that are offered as messianic and indispensable prove precisely the inevitable failure of inclusion.

At The Crossroads: Breaking with Identity

There is a common cultural machinery at work, and its tools are fearfully similar from one group to the next. It defines and stigmatizes them. It pushes them out to the margins. And then it flushes them out the bottom economically. You can't change that system by just getting your own rights, tinkering with the engine and leaving. You have to take on the whole machine.¹²

Serving as an analytic category, “identity” can erase the intersections of a differing multiplicity of social experiences. In privileging one identity in the struggle for recognition, the intersections of these differences are elided, invalidating those who are variably engaged by racial, ethnic, sexual, gendered, classed, and other kinds of differences (and the resulting oppression).

In fashioning a trans identity politics to address trans struggles—the concrete violence experienced on systematic and everyday levels—one may also consider the specific threat of the gatekeeping rhetoric that Trans Exclusionary Radical Feminists (so-called TERFs) and some Gender-Critical Feminists have employed to argue for an exclusive, essentialist category of womanhood against trans politics. The exclusionary policing of womanhood on a biological basis of sex¹³ neglects the intersections of oppression in the social domain, whereas, instead of being born into a fixed identity that intrinsically constitutes the self and everything thereafter, one

¹² Riki Wilchins, *Read My Lips: Sexual Subversion and the End of Gender*, 81.

¹³ It has been suggested in queer and poststructuralist feminism that there is not a “biological sex” to begin with as in the analogy of actors performing their gender on a predetermined sexed stage. Rather, our understanding of sex is inextricably intertwined with gender, and the processes through which these categories are written on our bodies also shape them as they become the very terms we use to understand ourselves in the matrix of intelligibility. In other words (that is, in Butlerian terminology), gender is performative, not performance, and in the performativity of gender thus resides the potentiality of variance and subversion (Butler, 1990).

becomes oneself as a result of social processes, a matrix of interrelations that are historically conditioned in instituting those processes as violent and oppressive. The process does not proceed singularly: patriarchy functions not by subjugating all women to all men, as, for instance, the identity-based ascription of “male privilege” as summarily inherent in men would neglect the fact that effeminate men and gender-variant people are similarly oppressed and subjugated, precisely because the oppression is not ontologically sexed but working within a complex schematism of power that initiates categories of normality, which simultaneously produces and enforces divisive identitarian categories as the “master’s tools” for basing and perpetuating oppression.

The notion of transmisogyny, coined by Julia Serano to describe the intersection of sexism and transphobia, sheds light on the plurality of mechanisms that work to produce gender subjugation. In *Whipping Girl: A Transsexual Woman on Sexism and the Scapegoating of Femininity*, Serano accounts for gender oppression from a transfeminine perspective. She divides sexism between oppositional sexism (the imperative that male and female remain distinct, “opposing”, and fixed) and traditional sexism (the denigration of women, or those associated with them, and the elevation of men and the manly over them). As trans women were subject to this in total force at once (often quite visibly defying normative imperatives around sex assigned at birth, and at once seeming to opt for womanhood over manhood), they are effectively intersectional victims of both sustaining structures that define sexism (Serano, 2007). This analysis of gendered oppression, by calling attention to the intersectional dynamics of oppressive forces, brings to the fore a transfeminist vision beyond the binarized conceptions of identities. It also makes room for questioning whether “woman” is a sufficient analytic category for accounting for the various forms of oppression that women can experience in a sexist society (Stryker & Bettcher, 2016).

The case of intersectional oppression then sheds light on the failure of an identity model of recognition that addresses misrecognition only from the circumstances giving rise to—and thus limited by—an ahistorical construct masked under the self-defeating language of originality.

Status-based Recognition and The Potentiality for Coalitional Resistance

An open coalition, then, will affirm identities that are alternately instituted and relinquished according to the purposes at hand; it will be an open assemblage that permits of multiple convergences and divergences without obedience to a normative telos of definitional closure.¹⁴

Departing from the insufficiency of an identity politics of difference, I argue for a framework of recognition as a question not of identity but social status.

¹⁴ Judith Butler, *Gender Trouble*, 22.

Status-based rather than identity-based interests should be useful to addressing common concerns to avoid homogenizing (gendered) groups of people. This way, the fundamental problem lies not on identitarian categories which posit a differential ontology of sex/gender, mobilized to police, for instance, the limits of womanhood or intelligible modalities of trans selfhood. Instead, it should describe and address the mutual processes of social becoming beyond identity models of recognition (Fraser, 2000; Saeidzadeh & Strid, 2020). One is not recognized in an ontological, self-affirming relation to a differing other; rather, it is through the shared relations of oppression, that group interests exist on the political ideal of not being oppressed, and it is through this shared history and relationality of status and needs that recognition is secured to create conditions for alterity. This then shifts the focus from identity-based self-realization to a picture of self-creation that aligns with the needs and desires of trans individuals without calling out to restrictive categories that function by the hierarchizing grammar of intelligibility. The trans subject is not formed passively in response to hetero-patriarchal systems but actively enacts a logic of self-fashioning that corresponds to the very concreteness of trans embodiment. Hormones and other gender-affirming medicine are not acquired as intrinsically quintessential of being trans, as a depathologized basis of choice and agency should be granted. In response to doubts and skepticism, trans medicine is not employed to reinforce stereotypical gender binary by self-reification—such a line of reasoning must fail when one considers the similarity instead of difference that is found in its false opposite, as the wish to “become woman” is an impossible endeavor that exists for cisgender women as well—the idea of womanhood as such is postulated as a semblance of impossible perfection pressured by a patriarchal world. The trans experience, on the other hand, immediately enacts a displacement of that order with variance and variability—it entails the potentiality for a mode of living that splits with the cissexual commonsense which has imposed its hegemonic logic on all gendered existences. The status of being trans and women and men, classed, raced, gendered, dis/abled, thus emerges from the same genealogy of oppression and resistance—a world that operates by prioritizing certain modes of existence while excluding others as unviable and misrecognized. As the identitarian model only conceptualizes the existing gender hierarchy on a unitary line, in which trans people are subordinated to cis people, women to men, it is the same world—precisely so—that subordinates working class men of color to white middle class women, effeminate men to masculine men, and so on and so forth, alongside an inexhaustible mosaic of convergence wherein the designation of clearly legible identities becomes counter-productive. To recognize this dynamic of shared statuses is thus to rise against the totality of social relations that produce varying forms of oppression, laying ground for a coalitional vision of the future of trans resistance and liberation.

Bibliography

- Bettcher, T. M. (2007). Evil Deceivers and Make-Believers: On Transphobic Violence and the Politics of Illusion. *Hypatia*, 22(3), 43–65. <https://doi.org/10.1111/j.1527-2001.2007.tb01090.x>
- Bettcher, T. M. (2014). Trapped in the Wrong Theory: Rethinking Trans Oppression and Resistance. *Signs: Journal of Women in Culture and Society*, 39(2), 383–406. <https://doi.org/10.1086/673088>
- Bornstein, K. (1994). *Gender Outlaw: On Men, Women and the Rest of Us*. Routledge.

- Butler, J. (1990). *Gender trouble: Feminism and the Subversion of Identity*. Routledge.
- Butler, J. (2005). *Giving an account of oneself*. New York: Fordham University Press.
- Butler, J. P. (2006). *The Psychic Life of Power: Theories in Subjection*. Stanford University Press.
- Fraser, N. (2000b). Rethinking recognition. *New Left Review*, 3, 107–120.
- Gleeson, J. J. (2020). Conflict, Recognition & Gender Transitions: Normativity Struggles in Contemporary Feminism's Trans-Inclusion Dispute. *Interfere*.
- Hausman, B. L. (1995). *Changing Sex: Transsexualism, Technology, And The Idea of Gender*. Duke University Press.
- Heyes, C. J., & Latham, J. R. (2018). Trans Surgeries and Cosmetic Surgeries. *TSQ: Transgender Studies Quarterly*, 5(2), 174–189. <https://doi.org/10.1215/23289252-4348617>
- Honneth, A. (1995). *The struggle for recognition: The moral grammar of social conflicts* (J. Anderson, Trans.). Cambridge: Polity Press.
- Knox, Liam. "Media's 'detransition' narrative is fueling misconceptions, trans advocates say." NBC News, 19 Dec 2019.
- Knudson, Gail, et al. (2016). "Position Statement on Medical Necessity of Treatment, Sex Reassignment, and Insurance Coverage in the U.S.A." World Professional Association for Transgender Health (WPATH), December 21. www.wpath.org/site_page.cfm?pk_association_webpage_menu=1352&pk_association_webpage=3947.
- Mason-Schrock, D. (1996). Transsexuals' Narrative Construction of the "True Self". *Social Psychology Quarterly*, 59(3), 176. <https://doi.org/10.2307/2787018>
- Plemons, E. (2017). *The Look of a Woman: Facial Feminization Surgery and the Aims of Trans-medicine*. Duke University Press.
- Plemons, E., & Straayer, C. (2018). Introduction: Reframing the Surgical. *TSQ: Transgender Studies Quarterly*, 5(2), 164–173. <https://doi.org/10.1215/23289252-4348605>
- Saeidzadeh, Z., & Strid, S. (2020). Trans* Politics and the Feminist Project: Revisiting the Politics of Recognition to Resolve Impasses. *Politics and Governance*, 8(3), 312–320. <https://doi.org/10.17645/pag.v8i3.2825>
- Salamon, G. (2010). *Assuming a body: Transgender and rhetorics of materiality*. New York: Columbia University Press.
- Schegan, N. (2011). Queering the Center by Centering the Queer. *Shifting Ground*, 111–144. <https://doi.org/10.1093/acprof:osobl/9780195395112.003.0008>
- Serano, J. (2007). *Whipping Girl*. Seal Press.
- Stone, S. (1991). *The "Empire" Strikes Back: A Posttranssexual Manifesto*. Austin, TX: University of Texas.
- Stryker, S. (2004). Transgender activism. *Glbq*. Retrieved from <http://www.glbqarchive.com>
- Stryker, S., & Bettcher, T. M. (2016). Introduction: Trans/feminisms. *Transgender Studies Quarterly*, 3(1/2), 5–14.
- Taylor, C. (1994). The Politics of Recognition. *Multiculturalism*, 25–74. <https://doi.org/10.2307/j.ctt7snkj.6>

Valentine, D. (2007). *Imagining Transgender: An Ethnography of a Category*.
Duke University Press.

Wilchins, R. A. (1997). *Read My Lips: Sexual Subversion and the End of Gender*.
Magnus Books.



An Examination of Descartes' Doubt and the Foundations of Certain Knowledge

Nathaniel Chapman

Author Background: *Nathaniel Chapman grew up in the United States and currently attends St. Paul's School, in Concord, New Hampshire in the United States. His Pioneer research concentration was in the field of philosophy and titled "Descartes' Meditations on First Philosophy."*

Abstract

Descartes' primary goal is that his philosophical system be immune to skepticism of any degree. Therefore, to uncover certain knowledge, this paper will evaluate his argument for the existence of the "I" against his own metric of indubitability. This essay consists of three sections. I begin in the first section with an exposition of Descartes' famous *Cogito argument* as depicted in the *Meditations on First Philosophy*. In the second section, I will examine two salient interpretations—the "performative" and the "deductive"—from prominent philosophers Jaako Hintikka and Harry Frankfurt, respectively. In the third and final section, I will present a holistic evaluation and critique of the doubt-resistance of the *Cogito*. My critique is twofold: that Descartes' method of doubt is somewhat insincere, and secondly, that the use of a truly rigorous skepticism against his epistemology would render almost all knowledge and information completely unintelligible.

1. Introduction

The 17th-century philosopher and polymath René Descartes (1596-1650), whose work is considered by many to be foundational to modern western thought, dedicated himself to the construction of a secure epistemic foundation for philosophical truth and the sciences.

The historical and intellectual context surrounding this pursuit helps to illustrate why Descartes pursued such a foundation. His time was characterized by rapid scientific development and a surplus of new knowledge claims and understandings of the world. However, frequent revolutions in understanding undermined the certainty of truth. As new theories supplanted the old, the wheels of science seemed to devour themselves. Descartes was a system builder, and all-around he observed cracks in the foundations of what was considered truth.

For this reason, he attempted to craft a system that would not be undermined by scientific discoveries: a foundation upon which philosophers and scientists could build indubitable knowledge for centuries to come. In Descartes' philosophical treatise *The Meditations on First Philosophy* (1641), he constructs this system through the perspective of an individual, the 'Thinker'. By doubting all that can be doubted, this Thinker goes on to purportedly prove the existence of God, the soul, and the external world. After successfully suspending his belief in the existence of nearly every concept and object in the first meditation, he comes to a certain existent in the second meditation: his existence, "So after considering everything very thoroughly, I must finally conclude that this proposition, I am, I exist, is necessarily true whenever it is put forward by me or conceived by mind" (AT VII 25). This argument is commonly known as the '*Cogito*,' and will be referred to as such throughout this paper. From the validity of this datum, his system stands or falls.

2. Exposition of Descartes' Methodology and the Initial Formulation of the *Cogito*

The *Meditations on First Philosophy* contains the most comprehensive illustration of his approach to constructing an epistemological system immune to skeptical doubt upon the certainty of the *Cogito*. Therefore, this text, particularly the first two meditations, one concerned with doubt and the other with establishing the *Cogito*—What Can be Called into Doubt and The Nature of the Human Mind, and How it is Better Known than the Body—are of primary importance in a thorough exposition of his argument for the existence of the "I".

In the First Meditation, the Thinker begins his methodical and sweeping elimination of all truths that can be doubted; this section serves as a good starting point for understanding the criteria by which the *Cogito* will be judged. At some point, after reflecting upon the numerous false beliefs that he accepted earlier in his life, he "realized that it was necessary... to demolish everything completely and start again right from the foundations if [he] wanted to establish anything at all in the sciences that was stable and likely to last" (AT VII 17). This illustrates that the Thinker's (and by extension, Descartes') primary goal in undergoing these meditations is the creation of a system that is immune to doubt and the changing paradigms of new knowledge. The Thinker goes on to state that, "Reason now leads me to think that I should hold back my assent from opinions which are not completely certain and indubitable just as carefully as I do from those that are patently false" (AT VII 18). This statement is central to the whole of Descartes' philosophical pursuit—the creation of a system of knowledge that is beyond doubt and uncertainty. Turning briefly to the *Cogito* that will come later in the Second Meditation, whether one's existence is truly beyond any doubt will determine the argument's success as a secure building block for knowledge.

Continuing with the First Meditation, the Thinker takes it upon himself to doubt his previously held beliefs. However, it would be an arduous and inefficient task to eliminate one by one all his prior notions. Consequently, he endeavors to undermine large swaths of beliefs by doubting the mechanism through which he gained them: the sense apparatus. His systematic process of doubt will be hereafter referred to as the 'Method'.

To undertake the Method, he utilizes three distinct arguments for skepticism, each more potent than the last: the occasional deception of the senses, the dream argument, and the malicious demon argument. The latter two are by far the most

powerful. Reflecting on the dream argument, the Thinker asserts, "as I think about this more carefully, I see plainly that there are never any sure signs by means of which being awake can be distinguished from being asleep" (AT VII 19). If this is accepted, then one cannot be sure of the truth-value of one's memories or current experience of their perceptions. Thus, empirical knowledge must be doubted, but some information remains unscathed. For example, in both our dreams and our waking lives, we perceive colors, numbers, and the extensions of physical objects in space. On this basis, knowledge concerning mathematics, and other more general abstractions, is still certain.

But this alone does not satisfy the Thinker's experiment in extreme skepticism, as he still questions, "how do I know that God has not brought it about that I too go wrong every time I add two and three or count the sides of a square." (AT VII 20) The Thinker then undermines his certainty in these beliefs by postulating the possibility that "some malicious demon of the utmost power and cunning has employed all his energies in order to deceive me" (AT VII 22). This demon undermines every perception and every piece of knowledge that the Thinker had previously accepted systematically. Furthermore, as this demon makes basic arithmetic fallible, it seems that this demon is even capable of impairing the Thinker's reason. In summary, the First Meditation elucidates two primary ideas concerning the *Cogito*: Descartes' motivation for the whole endeavor, and the essential criteria of indubitability. If Descartes is to be successful, then the *Cogito* must be immune to doubt even in the face of impaired and possibly manipulated reason.

After establishing his goals and methodology in the First Meditation, the Second Meditation illuminates the *Cogito* in its simplicity and purported certainty. The Thinker, unnerved by the impossibility of trusting the existence of the external world and his own body, searches his faculties for some process whose validity is not sabotaged by his malicious deceiver.

He believes he finds a haven in his general thought processes: "Thinking? At last I have discovered it - thought; this alone is inseparable from me. I am, I exist - that is certain. But for how long? For as long as I am thinking" (AT VII 27). The Thinker believes that, for the duration of time he has thoughts, he cannot doubt his own existence. At a first glance, this seems to be a logical inference, deriving the conclusion of the Thinker's existence from the evident necessity of thoughts having a prior cause—a thinker. However, the method of doubt calls even common-sense notions such as the implicit connection between thoughts and one who thinks into question. Aware of this syllogistic danger, Descartes clarifies the *Cogito* in the "Second Replies" (a collection of objections and replies compiled at the end of *The Meditations*), stating that when one formulates the *Cogito*, "He does not deduce existence from thought by means of a syllogism, but... as something self evident by a simple intuition of the mind" (AT VII 141). Descartes' *Cogito* is portrayed as an instantaneous intuition as opposed to a strictly logical inference. In other works, the well-known *cogito ergo sum* (an inference) can be found, but not here. This makes the *Cogito* out to be a powerful tool, as the evidence for its validity is automatically available whenever he turns his reason upon himself.

For the time being, the certainty of the *Cogito* even manages to survive the Thinker's hampered reason, as it does not rely on the *ergo* (therefore) but is instantaneously known whenever it is reflected upon. The validity of this formulation of the *Cogito*, then, seems to rest on the certainty of the Thinker's intuition. Whether or not the mechanism of intuition can hold up to extreme skeptical doubt remains to be seen. In the next section, Hintikka's and Frankfurt's interpretations illustrate two distinct manners in which the *Cogito* can be understood to be indubitable.

3. Hintikka's and Frankfurt's Interpretations

3.1 Hintikka's Performative Interpretation

Jaako Hintikka articulates the Performative Interpretation in his article "Cogito, Ergo Sum: Inference or Performance?" (1962). By performative, Hintikka means that it is the very act of uttering the *Cogito* that proves its validity. Over 300 hundred years since the life and work of Descartes, Hintikka reinterpreted definitions of thought and uncovered a novel proof of the *Cogito*. In its culmination, the Performative Interpretation proves itself by its negation; the utterance of the statement, "I think, but I do not exist" proves one's existence. This argument is intelligible if we understand Hintikka's distinct terminology and reasoning.

Most importantly, the proof of this interpretation relies on his notion of 'existential inconsistency', which "may be defined as follows: Let 'p' be a sentence and 'a' a singular term (e.g. a name, a pronoun, or a definite description). We shall say that p is existentially inconsistent for the person referred to by a to utter if and only if the longer sentence 'p; and 'a' exists' is inconsistent" (Hintikka, Inference or Performance? 10). This may be a difficult definition for non-logicians to grasp, but the idea can be made clear. Imagine that an individual (referred to as a) utters—or thinks—a statement (referred to as p). The statement p (I do not exist) is existentially inconsistent with the reality that the individual, a, is uttering p. This definition may still seem vague, however, following the argument to its conclusion elucidates its utility.

The first step in Hintikka's argument is properly conceiving thought and language. As stated earlier, Descartes' recognition of the absolute certainty of his existence came from the one faculty he refused to doubt: thought. But what is thought exactly? Hintikka states that thought, as Descartes used it, would have been more adequately defined as a form of persuasion, "Descartes' verbs *cogitare* (to think) and *dubitare* (to doubt) are not... the most accurate ones for describing the act through which the sentence, 'I don't exist' defeats itself... what he cannot do is persuade anybody (including himself) that he does not exist" (Hintikka, Inference or Performance? 18). According to Hintikka, thought is a kind of language, particularly, persuasive language. When we attempt to think through a statement of truth, we are attempting to persuade ourselves one way or another. If I am thinking through the statement, I exist, then I am attempting to persuade myself of one of two conclusions: I do exist or I do not exist. This follows because thought is only possible through the mode of, and derived from, language.

And what is the foundational purpose of language? Hintikka writes, "Normally a speaker wants his hearer to believe what he says. The whole 'language-game'... is based on the assumption that this is normally the case" (Hintikka, Inference or Performance? 12). Thus, if a speaker wishes to convince someone of something, their statements must be consistent with the conclusion they are advancing. If they tell someone that they do not exist, the act of that utterance is absurd. It is absurd because the utterance itself adds credibility to the notion of their existence in the other person's mind. "The pointlessness of existentially inconsistent statements is therefore due to the fact that they automatically destroy one of the major purposes which the act of uttering a declarative sentence normally has" (Hintikka, Inference or Performance? 12). The 'major purpose' mentioned is convincibility. Existentially inconsistent statements, in their utterance, significantly diminish the ability for one to accept their conclusion and aid in accepting the contrary.

And so, in relating this understanding of language and thought to the Thinker's utterance of the *Cogito*, Hintikka believes that the certainty Descartes saw in the statement, "I am, I exist", (AT VII 27), was actually found within the implicit impossibility of persuading oneself of its negation, "I do not exist". Whenever the Thinker attempts to doubt his existence (persuade himself of the lack of his being), he perceives the certainty of his existence. It is impossible for him to persuade himself of his own negation because the very process of doing so asserts the opposite. Understood in this manner, the *Cogito* argument is indubitable, as in it is literally impossible to doubt, or to persuade oneself, of its negation.

3.2 Frankfurt's Deductive Interpretation

Henry G. Frankfurt simultaneously launches a critique of Hintikka's understanding while putting forward his own interpretation of Descartes' *Cogito* in his essay, "Descartes's Discussion of his Existence in the Second Mediation" (1966). Frankfurt finds Hintikka's focus on the performative character of the *Cogito* misdirected and asserts that the strength of Descartes' argument lies not in its performative character, but in a deductive proof. In his response to Hintikka, Frankfurt states, "Hintikka's emphasis on performativeness is, in my opinion, not so much erroneous as inadequate... He does not tell the whole story and, because of his logical and textual errors, he does not even tell his part of the story well" (Frankfurt, Discussion of the *Cogito*, 4). Frankfurt does not deny that there is a performative character to this proof of existence. However, he disagrees with Hintikka that the indubitable nature of the *Cogito* is fundamentally derived from its performative nature.

Instead, Frankfurt focuses on the fact that from the Thinker's perspective, the malicious demon is the most convincing reason to doubt one's convictions. Therefore, if we are to find truth, we must prove that we can have knowledge whilst under such manipulation. Following this, Frankfurt illustrates the evidence of the Thinker's existence by tracing the way he could be deceived. "I am deceived in thinking that I exist. If I am deceived about anything, then I exist. Hence, if I am deceived in thinking that I exist, then I do exist. Therefore, I do exist. If I do exist, then I am not deceived in thinking that I exist. Therefore, I am not deceived in thinking that I exist" (Frankfurt, Discussion of the *Cogito*, 13). From this line of argumentation, it is not that the *Cogito* is proved by its utterance as Hintikka understood it, but that under the method of doubt used, it is logically impossible for one to be deceived about one's existence.

Whenever one considers the question of one's existence along with the mechanics of the Method, the doubt destroys itself. "This conclusion in no way denies, of course, that *sum* (I am) is logically contingent. When Descartes describes *sum* as necessarily true on certain occasions, he does not mean that on those occasions it possesses the logical property of necessary truth. He means that it is a logically necessary truth that *sum* is true on those occasions" (Frankfurt, Discussion of the *Cogito*, 15). Thus, Frankfurt asserts, that the primary indubitability of the *Cogito* is found not in its performative aspect, but its logical undeniability. Both Frankfurt and Hintikka's understandings provide the indubitable nature of the *Cogito* whenever it is considered. The crucial difference being that the affirmation of one's existence in Hintikka results from the performative nature of thought, while in Frankfurt it results from the logical impossibility of doubting under the influence of the malicious demon of Descartes' Method.

4. The Critique from Passive Skepticism

After considering the argument as presented by Descartes, Hintikka's performative understanding, and Frankfurt's proof by deduction it seems evident that the *Cogito* is resistant to doubt in two ways. Firstly, it is indubitable to the degree that it is impossible to persuade oneself of its negation, and secondly, because the existence of a manipulative agent makes the existence of self a strict logical necessity. These understandings of the *Cogito* are viable—but only under what may be defined as an 'active skepticism'.

This notion of active skepticism is referenced by Hintikka briefly in his characterization of Descartes' Method, stating it "does not consist in the giving up of all opinions, as a skeptic's doubt might... It amounts to an active attempt to think the contrary of what we usually believe" (Hintikka, *Inference or Performance?* 17). Therefore, we may define active skepticism as postulating possible (even if improbable) reasons—the dream argument, the malicious demon argument—for doubting what we hold to be true. This is the type of doubt that is contained within the Method. This is problematic because Descartes' declared aim was to provide something immune to all forms of doubt. For him to do so, his foundation must attain immunity from every form of skepticism.

As well Hintikka implied that there is a difference between the active skepticism used, and the doubt of an 'actual skeptic'. The doubt of the actual skeptic may pose a dangerous challenge to Descartes' *Cogito*. This second type of doubt may be defined as a 'passive skepticism'. That is, a fundamental refusal to hold any assumptions at all, without positing any positive reasons to doubt. Using passive skepticism knowledge and the reliability of all thinking in Descartes' argument seems to crumble.

Firstly, Descartes seems to rely on the Principle of Sufficient Reason (PSR). Although the PSR had not been formulated in Descartes' time, the inference from the experience of thoughts, desires, and images, to the Thinker being a "thing that thinks" (AT VII 28), closely resembles the idea that for every positive fact, there must be a sufficient cause (The Stanford Encyclopedia of Philosophy 2021). Not once does the Thinker ever question the idea that for there to be thoughts, there must be causes, i.e., thinkers. While there is an incredibly nuanced debate surrounding the validity of the PSR, under the requirements of passive skepticism, it would be denied. Therefore, there is no reason to believe that there is any connection between events and causes. A passive skeptic would not know anything more than the experiences of thoughts as they seem to exist in the present moment.

Secondly, and most damaging to Descartes' system, a strict passive skepticism requires the withdrawal of our belief in the assumptions of language and definition. Early in the Meditations, the Thinker recognizes that the dream and malicious demon argument make memories fallible (AT VII 19). However, the Thinker does not pursue significant implications of this fallibility. If memories cannot be trusted, then the very definition of the words used in philosophical argumentation (or any other linguistic endeavor) cannot be relied on to be consistent.

Language is gained through a developmental process. The meanings of the words we use are always dependent on our memory of the experiences that formed our vocabulary. For example, after we are born, we develop language through socialization and growth in our intellect and other faculties. If we deny the certainty of our memories, then we must deny the certainty of the meaning of our language. Under both passive skepticism (there is no reason to trust memories) and active skepticism (memories can be easily manipulated by the malicious demon), our

memories of development, and thus our language, are called into doubt. Both the active and the passive skeptic easily doubt Descartes' system, provided they are willing to sacrifice language. However, sacrificing language amounts to completely disavowing rational inquiry, as one cannot reason if the words they use to do so are untrustworthy. The consequences of this skeptical conclusion are dire: reasoning becomes impossible. Without consistent definitions over time, existence is rendered unintelligible. Every formulation of information is eternally called into doubt. Here, Descartes falls short of his criterion of indubitability. From the skeptic's perspective, not only does the *Cogito* fail, but all knowledge—save strictly parameterized information referring to the images of immediate experience—is made utterly dubitable.

5. Conclusion

Descartes sought a system that would be immune to doubt and revolutions in scientific and philosophical thought. Through the exposition of the argument, Hintikka's and Frankfurt's perspectives, and the passive skepticism attack, whether Descartes succeeded depends on the degree to which one is willing to take their skepticism. Clearly, if passive skepticism as it is articulated in this paper is to be accepted, then Descartes fails on all fronts. Under passive skepticism, the *Cogito* is flimsy, his epistemic foundation of cement turns to sand, and there is little hope of any knowledge in the world.

But should we accept passive skepticism? That question is beyond the scope of this paper. However, from a pragmatic perspective, it does seem self-evident that there is zero use in a perspective that refuses to make any assumptions at all. When the Method is not stretched to its absolute extreme (the meaning of words is still reliable), there does in fact seem to be significant justification for the indubitability of the *Cogito*. This is true along both the performative and logically inferred interpretations presented by Hintikka and Frankfurt.

In conclusion, one only needs to look at Descartes' continued influence in the worlds of philosophy and even pop culture to answer the question of whether his philosophy has endured. As for the *Cogito*, if there is to be knowledge at all, it is generally beyond doubt. But equally so, if one is willing to accept a world completely without reason and language, then the *Cogito* may be doubted. Clearly, if we aim to progress in the attainment of knowledge and truth, it is necessary to accept the former for now.

Works Cited

- Descartes, René, and John Cottingham. *Meditations on First Philosophy: With Selections from the Objections and Replies*. Cambridge (Cambridgeshire): Cambridge University Press, 1986.
- Hintikka, Jaakko. "Cogito, Ergo Sum: Inference or Performance?" *The Philosophical Review*, vol. 71, no. 1, [Duke University Press, *Philosophical Review*], 1962, pp. 3–32, <https://doi.org/10.2307/2183678>.
- Frankfurt, Harry G. "Descartes's Discussion of His Existence in the Second Meditation." *The Philosophical Review*, vol. 75, no. 3, [Duke University Press, *Philosophical Review*], 1966, pp. 329–56, <https://doi.org/10.2307/2183145>.

Dika, Tarek R., "Descartes' Method", The Stanford Encyclopedia of Philosophy (Summer 2020 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2020/entries/descartes-method/>.

Melamed, Yitzhak Y. and Martin Lin, "Principle of Sufficient Reason," The Stanford Encyclopedia of Philosophy (Summer 2021 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2021/entries/sufficient-reason/>



The Relationship Between Religious Practice and Depression: Intrinsic Religiousness as Moderator

Qianyu Liang

Author Background: *Qianyu Liang grew up in China and currently attends Shenzhen College of International Education in Shenzhen, China. Her Pioneer research concentration was in the field of psychology and titled "Stress and Health."*

Abstract

Previous research suggested that religious practices may help reduce the degree of depression and improve an individual's mental health condition (e.g., Worthington & Kuru, 1996; Funder, 2002; Joiner, Perez, & Walker, 2002). This study examined two primary research questions: Firstly, what is the relationship between religious practice and depression? Secondly, does the relationship between religious practices and depression depend on intrinsic religiousness? Self-report data from 212 female participants from various ethnic and educational backgrounds were collected via questionnaires with Duke University Religion Index (DUREL) and Beck Depression Inventory (BDI). Semi-structured interviews with five female respondents were also conducted to collect qualitative data and shed light on the quantitative data analysis. SPSS PROCESS MACRO v4.0 was used to test the hypothesis. The results suggest that the levels of intrinsic religiousness (IR) moderated the relationship between organizational religious activity (ORA) and depression (BDI), and the levels of intrinsic religiousness (IR) also moderated the relationship between non-organizational religious activity (NORA) and depression (BDI). Implications of these results for future studies on religious practices and religiosity-based therapeutic criteria for depressive symptoms were discussed.

Keywords: religion; mental health; intrinsic religiousness (IR); organizational religious activity (ORA); non-organizational religious activity (NORA); Beck Depression Inventory (BDI); Duke University Religion Index (DUREL).

Introduction

Religious belief and practice have enduring prevalence in western societies. Recent polls concurrently indicate that more than 90% of US citizens believe in God or some kind of higher power, and 63% of them expressed that religion is "very important" in their lives (Esquivel, 2021). In European countries, religiousness seemingly declined in recent decades, but the personal belief in God remained very popular (Grosby,

2021). Observing this phenomenon, psychologists have long been interested in religiousness's influence on individuals' physical and psychological health conditions. Although historically disapproved by prominent psychologists, the current consensus among psychology academia states that certain forms of religiousness can consistently facilitate mental health (Graham, 2021). Intrinsic religiousness, in particular, is frequently discovered in research to be associated with better mental health. These discoveries call for a deeper inquiry into the subject: specifically, what moderates or explains the observed association between religiousness and mental health? This study will examine intrinsic religiousness as a possible moderator between individuals' religious activities and mental health conditions.

Before discussing the relevant studies, the religion-related variables examined in this study are first defined here. Organizational religious activity (ORA) refers to the attendance of public religious activities or services, including Bible study groups, meditation groups, and prayer groups. Non-organizational religious activity (NORA) refers to private religious activities, such as prayer, meditation, reading Bible, watching religious TV programs, or listening to religious broadcasts. Intrinsic religiousness (IR) refers to the degree of personal religious commitment and the extent to which religiousness is considered by the individuals as a crucial portion of their everyday lives or a source of meaning, motivation, and belief (Koenig & Büssing, 2010).

Distinguishing Between Intrinsic and Extrinsic Religiousness

Most previous studies on religion and mental health employ very simplistic measures of religiousness, sometimes surveying only a single question, such as how often subjects attend religious services or how firmly they observe the religious doctrines (Levin & Markides, 1988). Unsurprisingly, these poor research methods only give them inconclusive results (McCullough & Larson, 1999). More recent studies have become more sophisticated and developed their methods to include validated, multidimensional measures of religiosity. Fehring et al. (1997) used such a measure of religiosity, called the Intrinsic/Extrinsic Religiousness Scale, and discovered a strong negative relationship between cancer patients' religiousness and mental health, despite many scholars' premise that religion should provide relief and support (Gorsuch & McPherson, 1989).

Fehring et al. (1997) distinguished "intrinsic" and "extrinsic" religiousness, with the former tapping a combination of religious practices associated with the more spiritual aspect of religion and the latter corresponding closely to the rituals and practices associated with institutional religions. Of interest, they found a significant negative association between intrinsic religiousness and depression ($r = -.44$) and no relationship between extrinsic religiousness and depression ($r = .09$). Moreover, several other studies employing the distinction of "intrinsic" and "extrinsic" religion have obtained similar results (Fehring, Miller, & Shaw, 1997).

This proves that we must clarify the relationship between spirituality and institutional religion before we go on inquiring about the relationship between psychological well-being and spirituality. Spirituality can exist with or without a religious framework. In fact, the group of individuals who reject any particular religion may as well be ones who consider themselves very spiritual (Vaughan, Wittine, & Walsh, 1988). On the other hand, the term "religion" always denotes an organized system of doctrines, practices, and rituals (Emblen, 1992). Although religion may be a channel to direct spiritual expression, it doesn't provide spirituality readily on its own end. For example, some adherents of religion may rigidly follow

the traditional, social, and ritual aspects of the religion but largely ignore its spiritual aspects (Elkins, Hedstrom, Hughes, Leaf, & Saunders, 1988).

Relationship Between Religious Practices and Mental Health

Cross-sectional studies suggest that attending regular church services is associated with a reduced risk of PTSD, alcohol use disorders, lifetime depressive illness, and current suicidal ideation (Balbuena, Baetz, & Bowen, 2013). Smith et al. (2003) discovered a weak negative relationship between religiousness and depression in a meta-analysis of 147 studies (Smith, McCullough, & Poll, 2003). AbdAleati et al. (2014) systematically reviewed multiple faith groups and reported that the groups with higher scores on religiousness or religious activity described lower odds of depression, anxiety disorders, suicidal behaviors, and substance misuse problems (AbdAleati, Mohd Zaharim, & Mydin, 2014).

During the past few decades, as a large number of studies have been done, some clarity in the relationship between religiousness and health has begun to emerge. The most solid finding is that religious zeal in adolescents and adults tends to be positively associated with positive psychological health (e.g., greater subjective happiness, life accomplishment, and self-respect), social behavior (e.g., greater marital stability, lower rates of crime, delinquency, and suicide), and physical health (e.g., lower rates of illnesses such as coronary heart disease and cancer; greater resilience to deal with and recover from illness; and greater life expectancy) (Donahue, 1985; Desrosiers & Miller, 2007; Allen & Wang, 2014; Barton, Miller, Wickramaratne, Gameroff, & Weissman, 2013; Oman & Lukoff, 2018).

Factors associated with religiousness that influence symptoms of depression have received the most attention in the recent empirical literature. Many researchers have suggested that religiousness may reduce vulnerability to depressive symptoms through various substantive social mechanisms (Koenig, 1992).

The Present Study

A number of previous researches have suggested a correlational relationship between religious practices, including organizational ones and non-organizational ones, and decreased levels of depressive symptoms (Miller et al., 2012). Moreover, high levels of intrinsic religiousness have been discovered to be associated with greater improvement of depressive symptoms, fewer suicide attempts, and higher living qualities in a sample of patients diagnosed with major depressive disorders (Mosqueiro, da Rocha, & Fleck, 2015). Nevertheless, only a few previous pieces of research have examined the role of intrinsic religiousness as a moderator of the correlation between religious practices and depressive symptoms. In light of this, the present analysis aimed to evaluate the role of intrinsic religiousness in moderating the relationship between religious practices and the levels of depression among individuals.

Methods

Participants

The samples were recruited from the United States. These individuals were offered payment for their participation. These participants represent a wide range of age, race, and level of education. In all, 212 participants consented to participate. Fifty-five of the individuals left parts of the self-report questionnaires uncompleted or failed to obey the researchers' instructions. Thus, their data were removed from the quantitative phase of statistical analysis.

In total, there were 157 participants who had provided valid data for the study. The mean age of the overall sample was 32.86 years, and the standard deviation was 5.62. The minimum age is 22 years, and the maximum age is 48 years. All participants in the sample were female. Detailed demographic percentages can be found in Table 1 and Table 2 below.

Table 1. *Final Race Identified by Participants*

Final Race Identified by Participants	N	%
American Indian or Alaska Native Only	1	0.51%
Asian or Asian American Only	17	8.04%
Black or African American Only	4	1.92%
White or Caucasian Only	58	27.47%
More than one race (non-Latinx)	10	4.66%
Latinx Only	18	8.49%
More than one race (Latinx)	49	23.08%
Missing	55	25.93%
Total	212	100.00%

Table 2. *The Highest Level of Education Completed by Participants*

Education Level	N	%
High School	12	5.72%
Some College	35	16.54%
Community College	15	7.13%
Bachelor's Degree	47	22.27%
Graduate Degree	41	19.30%
None of the Above	8	3.76%
Missing	54	25.48%
Total	212	100.00%

Measures

Self-report Demographics

After giving informed consent, all the participants were required to complete a series of self-report questions on demographic information involving their age, sex, and educational degree. The questions include "What is the race that you are identified as?" and "What is the highest level of education that you have completed?"

Religiousness Measure

Religiosity and spirituality were assessed using the Duke University Religion Index (DUREL) English version. The DUREL is a 5-item Likert scale assessment tool with three dimensions of religiosity: organizational religious activity (ORA), non-organizational religious activity (NORA), and intrinsic religiosity (IR). The first item evaluates the dimension of ORA, which refers to the attendance of public religious activities or services, including Bible study groups, meditation groups, and prayer groups. The second item evaluates NORA, which refers to private religious activities, such as prayer, meditation, reading Bible, watching religious TV programs, or listening to religious broadcasts. The final three items evaluate the dimensions of IR, which involve the degree of personal religious commitment and the extent to which religiousness is considered by the individuals as a crucial portion of their everyday lives or a source of meaning, motivation, and belief. Here, intrinsic religiousness (IR) was a comparison to extrinsic religiousness (ER), which is used for other more important purposes, such as social mobility or financial success, rather than merely for religion's sake. IR, on the other hand, includes the pursuit of religion as an ultimate end in itself (Koenig & Büssing, 2010).

The DUREL had a range of scores from 5 to 27. It consists of three subscales to measure three aspects of religiousness. Each subscale evaluates one of these aspects. Subscale 1 is the first question in the DUREL, asking about the frequency of

attendance at organizational religious activities (ORA). Subscale 2 is the second question, asking about the frequency of non-organizational religious activities (NORA). Subscale 3 includes the final three items, measuring the levels of intrinsic religiousness (IR) (Koenig & Büssing, 2010).

In this study, organizational religious activity, non-organizational religious activity, and depression level were the three factors subjected to correlational analysis, while the level of intrinsic religiousness was considered as the moderator of the correlational relationship. ORA, NORA, and IR were measured by the DUREL.

Depression Measure

The depression level was measured by the Beck Depression Inventory (BDI). The BDI consists of 21 items, and each item is a 4-point Likert scale. Respondents are asked to rate each item based on four response choices according to their feelings, behavior patterns, and symptoms during the past few weeks.

The following guidelines for BDI cut-off scores had been established by the Center for Cognitive Therapy: scores from 0 through 9 indicate no to minimal depression; scores from 10 through 18 indicate mild to moderate depression; scores from 19 through 29 indicate moderate to severe depression; scores from 30 through 63 indicate severe depression (Beck & Steer, 1988).

Participants' degree of depression was operationalized by their scores on BDI in the present study.

Qualitative Interview

The present study used an explanatory sequential design, in which a larger quantitative phase was followed by a smaller qualitative phase, aimed at explaining the quantitative results. Therefore, apart from measuring the quantitative variables, qualitative interviews were also conducted in order to collect qualitative data and shed light on the analysis of the quantitative data.

In total, five semi-structured interviews were conducted with five respondents. The respondents came from Shenzhen, Hong Kong, and Los Angeles, and they were recruited via opportunity sampling. Their mean age was 23.62 years, and the standard deviation was 12.37. The minimum age was 16 years, and the maximum age was 48 years. All participants in the sample were Asian or Asian American female Christians.

The semi-structured interviews consisted of a series of open-ended questions to obtain each individual's experience of religious practices and depression. The interviews with the five respondents were based on the same set of standardized questions. Yet, some extra "probe" questions were added according to the respondent's answers to the fixed list of questions, which enabled the respondents to develop, clarify, or give examples of particular points — this added depth and detail to the qualitative responses.

The interviews with the respondents were recorded with their informed consent and transcribed into scripts for further analysis and comparison.

Results

Descriptive Statistics

Table 3 below lists the mean and standard deviation of the measured variables. Descriptive statistics of the participant's scores on the various measures indicated that this sample of individuals was considered normal in terms of the level of depression ($M = 8.51$, $SD = 8.41$), attended organizational religious activities ($M = 2.67$, $SD = 2.67$) slightly more frequently than non-organizational religious activities ($M = 2.37$, $SD = 1.80$), and were highly intrinsically religious ($M = 10.22$, $SD = 3.44$).

Independent sample t-tests were conducted to examine whether there was a significant difference of the measure variables between white participants and non-white participants. The results of the t-tests suggest that there was a significant difference in the highest level of education received by participants, $t(156) = 4.13$, $p < .001$, which indicated that white participants ($M = 4.17$, $SD = 1.08$) received higher levels of education than non-white participants ($M = 3.26$, $SD = 1.47$). There was a significant difference in the scores on the Beck Depression Inventory of white participants ($M = 5.84$, $SD = 5.95$) and non-white participants ($M = 10.26$, $SD = 9.31$); $t(122) = -2.95$, $p = 0.002$, which indicated that the non-white participants were more depressive than the white participants. The differences in organizational religious activity ($t(124) = -1.51$, $p = .134$), non-organizational religious activity ($t(124) = -1.47$, $p = .143$), and intrinsic religiousness ($t(124) = -1.49$, $p = .138$) were insignificant.

Table 3. *Descriptive Data*

	N	Minimum	Minimum	Mean	Std. Deviation
<i>Education level</i>	158	1.00	6.00	3.59	1.41
<i>Depressive symptoms</i>	124	.00	41.00	8.51	8.41
<i>Organizational religious activity</i>	126	1.00	6.00	2.67	1.54
<i>Non-organizational religious activity</i>	126	1.00	6.00	2.37	1.70
<i>Intrinsic religiousness</i>	126	3.00	15.00	10.22	3.45
<i>Valid N (listwise)</i>	123				

Table 4. Group Descriptives

	Final race identified by participants	N	Mean	Std. Deviation
<i>Education level</i> ***	White	58	4.17	1.08
	Non-white	100	3.26	1.47
<i>Depressive symptoms</i> *	White	49	5.84	5.95
	Non-white	75	10.26	9.31
<i>Organizational religious activity</i>	White	50	2.42	1.51
	Non-white	76	2.84	1.55
<i>Non-organizational religious activity</i>	White	50	2.10	1.58
	Non-white	76	2.55	1.75
<i>Intrinsic religiousness</i>	White	50	9.66	3.61
	Non-white	76	10.59	3.31

Note: The significant tests were two-tailed.

* $p < .05$, ** $p < .01$, *** $p < .001$

Intercorrelations of Variables

The bivariate Pearson's correlation coefficients between the participant's score on the Beck Depression Inventory (BDI), organizational religious activity (ORA), non-organizational religious activity (NORA), and intrinsic religiousness (IR) were listed in Table 5 below. See the table for all correlations.

Table 5. Correlation Table

Final Race Identified by Participants: White					
		Organizational Religious Activity	Non-organizational Religious Activity	Intrinsic Religiousness	Depressive Symptoms
Organizational Religious Activity	Pearson Correlation	1.00	.79**	.64**	.05
	Sig. (2-tailed)		<.001	<.001	.976
Non-organizational Religious Activity	Pearson Correlation	.79**	1.00	.64**	.00
	Sig. (2-tailed)	<.001		<.001	.023
Intrinsic religiousness	Pearson Correlation	.64**	.64**	1.00	.02
	Sig. (2-tailed)	<.001	<.001		.876

Final Race Identified by Participants: White					
		Organizational Religious Activity	Non-organizational Religious Activity	Intrinsic Religiousness	Depressive Symptoms
Depressive symptoms	Pearson Correlation	.05	.00	.02	1.00
	Sig. (2-tailed)	.716	.976	.876	
Final Race Identified by Participants: Non-White					
		Organizational Religious Activity	Non-organizational Religious Activity	Intrinsic Religiousness	Depressive Symptoms
Organizational Religious Activity	Pearson Correlation	1.00	.63**	.64**	.04
	Sig. (2-tailed)		<.001	<.001	.759
Non-organizational Religious Activity	Pearson Correlation	.63**	1.00	.50**	-.07
	Sig. (2-tailed)	<.001		<.001	.573
Intrinsic religiousness	Pearson Correlation	.64**	.50**	1.00	-.09
	Sig. (2-tailed)	<.001	<.001		.573
Depressive symptoms	Pearson Correlation	.04	-.07	-.09	1.00
	Sig. (2-tailed)	.769	.573	.424	

Notes:

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

*** Correlation is significant at the 0.001 level (2-tailed).

Among the white samples, there was a strong positive correlation between the scores on ORA and NORA, $r(48) = .79, p < .001$. The scores on NORA and IR were also positively correlated, $r(48) = .64, p < .001$, and this positive correlation as strong as the correlation between ORA and IR, $r(48) = .64, p < .001$. There was no significant correlation between the measured religious variables and the score on the Beck Depression Inventory. The intercorrelations seem similar among the non-white samples.

Hypothesis Testing

A regression analysis was conducted to test whether the components of religious activity predicted depression levels. Depression was set as the dependent variable. Race and education were controlled. The three predictors were organizational religious activities (ORA), non-organizational religious activities (NORA), and intrinsic religiousness (IR). The overall model was significant $F(5, 117) = 2.37$, $p = .044$, which suggests the main effect.

Table 6. ANOVA^a

	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	591.68	2	295.84	4.45	.014 ^b
	Residual	7987.00	120	66.56		
	Total	8578.68	122			
2	Regression	788.30	5	157.66	2.37	.044 ^c
	Residual	7790.38	117	66.58		
	Total	8578.68	122			

Notes:

a. Dependent Variable: *BDI_total*

b. Predictors: (Constant), Final race identified by the participant, What is the highest level of education that you have completed?

c. Predictors: (Constant), Final race identified by the participant, What is the highest level of education that you have completed?, non-organizational religious activity, intrinsic motivation, organizational religious activity

Next, SPSS PROCESS MACRO v4.0 was used to test whether the association between organizational religious activities (ORA) and the depressive symptoms (BDI) depended on intrinsic religiousness (IR). BDI was set as the dependent variable, ORA was set as the independent variable, and IR was set as the moderator. All models were controlled for education and race. The results suggest that the overall model was significant, $F(6, 166) = 2.72$, $p = .017$. The interaction term was significant, $B = -.35$, $SE = .17$, $p = .044$, suggesting that the influence of ORA on BDI did depend on the levels of IR.

Conditional effects analyses suggested that the relationship between ORA and BDI was significant at low and medium levels of IR. At low levels of IR, the relationship between ORA and BDI was $B = 3.16$, $SE = 1.22$, $p = .011$. At medium levels of IR, the relationship between ORA and BDI was $B = 1.77$, $SE = .80$, $p = .028$. At high levels of IR, the relationship between ORA and BDI was not significant, $B = .33$, $SE = .87$, $p = .070$.

Table 7. Conditional Effect Analysis

IR	Effect	se	t	p
-4.19	3.16	1.22	2.60	.011
-.19	1.77	.80	2.22	.028
3.97	.33	.86	.38	.702

Note: Score on depression (BDI) was set as the dependent variable, organizational religious activity (ORA) was set as the independent variable, and intrinsic religiousness (IR) was set as the moderator variable. All models were controlled for education and race.

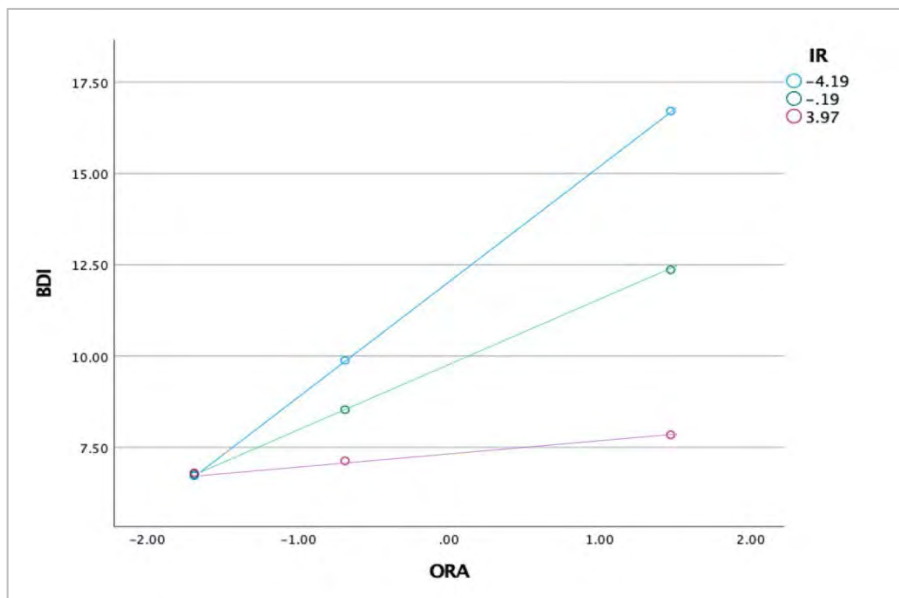


Figure 1. Conditional Effect Analysis: Intrinsic religiousness (IR) as a moderator of the relationship between organizational religious activity (ORA) and depression (BDI).

Besides, SPSS PROCESS MACRO v4.0 was also used to test whether the association between non-organizational religious activities (NORA) and the depressive symptoms (BDI) depended on intrinsic religiousness (IR). BDI was set as the dependent variable, NORA was set as the independent variable, and IR was set as the moderator variable. All models were controlled for education and race. The results suggest that the overall model was significant, $F(6, 166) = 2.72, p = .017$. The interaction term was significant, $B = -.34, SE = .17, p = .045$, suggesting that the influence of NORA on BDI did depend on the levels of IR.

Conditional effects analyses suggested that the relationship between NORA and BDI was not significant at low, medium, or high levels of IR. At low levels of IR, the relationship between NORA and BDI was $B = 1.41, SE = 1.15, p = .225$. At medium levels of IR, the relationship between NORA and BDI was $B = .46, SE = .68, p = .946$.

At high levels of IR, the relationship between ORA and BDI was $B = -1.37$, $SE = .73$, $p = .062$.

Table 8. *Conditional Effect Analysis*

IR	Effect	se	t	p
-4.19	1.41	1.15	1.22	.225
-.19	.046	.68	.07	.946
3.97	-1.37	.73	-1.89	.062

Note: Score on depression (BDI) was set as the dependent variable, non-organizational religious activity (NORA) was set as the independent variable, and intrinsic religiousness (IR) was set as the moderator variable. All models were controlled for education and race.

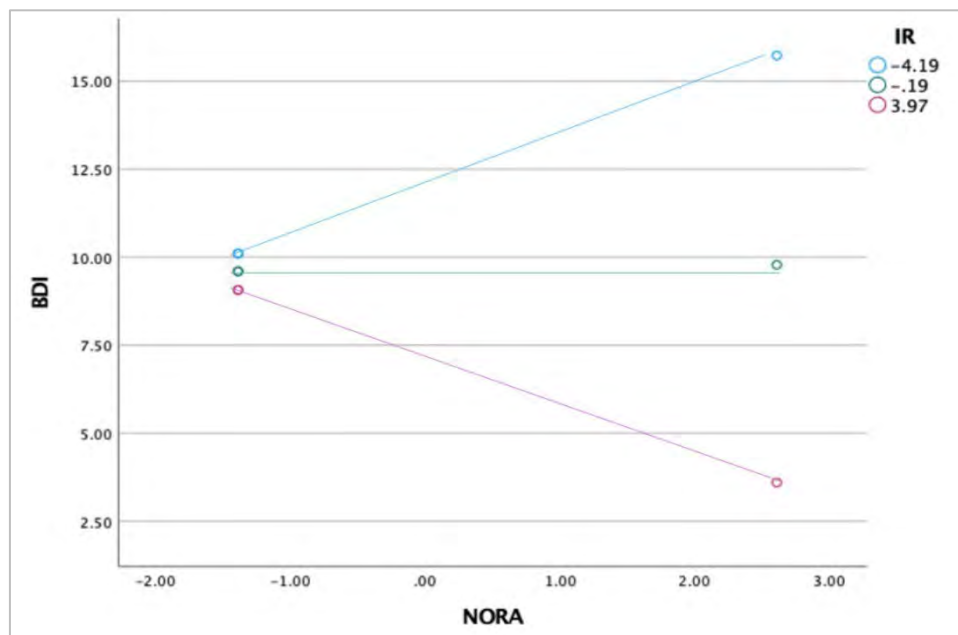


Figure 2. *Conditional Effect Analysis: Intrinsic religiousness (IR) as a moderator of the relationship between non-organizational religious activity (NORA) and depression (BDI).*

Explaining the Results

According to the regression analyses of the relationship between ORA, BDI, and IR, the scores on depressive symptoms increased with increased organizational religious activities. Considering this, a possible explanation for this positive correlation was that individuals who suffer from severe depressive symptoms tended to seek social support by attending public religious practices, such as church services or group-related spiritual activities, more than individuals who suffer from milder depressive symptoms. The fact that severe depressive symptoms propelled sufferers to religious

or spiritual practices could be evidenced by the participants' comments in their semi-structured interviews:

I felt that I was weighed down by the world in those years. I couldn't bear the feeling of helplessness on my own, so I went to the church to talk to the pastor almost every day and attend the spiritual healing group in the local community. (Participant 1)

It was my major depression disorder that brought me to the embrace of God. (Participant 2)

Back then, I couldn't fall asleep all night, and I had lost weight by more than 10 pounds. The experience of depression was so painful and despaired that I ultimately decided to beg for God's salvation by attending religious services at the local church every week. (Participant 3)

Moreover, the results suggested that the strength of the positive correlations between organizational religion and depression increased as the participants' levels of intrinsic religiousness decreased. The influence of intrinsic religiousness on the relationship between organizational religious activity and depression could be explained as follows: at lower intrinsic religiousness levels, individuals' religious faith was not effective enough to relieve themselves from depressive symptoms, and hence the scores on depression were relatively high at a certain level of organizational religious activities compared with the highly intrinsically religious individuals; at higher intrinsic religiousness levels, individuals' religious faith was relatively more effective in relieving the depressive symptoms, and hence the scores on depression were comparatively lower at a certain level of organizational religious activities. In light of this, intrinsic religiousness could be considered as a moderator in the correlation between organizational religious activities and depression.

Besides, regression analyses of the relationship between NORA, BDI, and IR suggested that at lower levels of intrinsic religiousness, the scores on depressive symptoms increased as non-organizational religious activities increased; at medium levels of intrinsic religiousness, non-organizational religious activities almost did not affect the levels of depression; at higher levels of intrinsic religiousness, the scores on depressive symptoms decreased with the increase in non-organizational religious activities. These results indicated that only when individuals' intrinsic religiousness was high could the non-organizational religious practices help reduce depression levels. The respondents' comments on their own experiences of depression were consistent with the results:

At that time, I felt that God was the ultimate meaning of my life. My faith in the belief that God loved me supported me in overcoming my physical ailments and emotional difficulties. Every time I prayed to God silently, I could feel the warmth and energy endowed by God in my heart, and those feelings indeed helped me a lot in my recovery from depression. (Participant 2)

I believe that it was my unshakable faith in God that saved me from those years of depression. (Participant 4)

In light of this, intrinsic religiousness could also be regarded as a moderator in the correlation between non-organizational religious activities and depressive symptoms.

Table 9. *Qualitative Interviews*

Topic	Theme	Representative Quotes
Religious or Spiritual Activities	A source of mental health support	<p>I felt that I was weighed down by the world in those years. I couldn't bear the feeling of helplessness on my own, so I went to the church to talk to the pastor almost every day and attend the spiritual healing group in the local community. (Participant 1)</p> <p>It was my major depression disorder that brought me to the embrace of God. (Participant 2)</p> <p>Back then, I couldn't fall asleep all night, and I had lost weight by more than 10 pounds. The experience of depression was so painful and despaired that I ultimately decided to beg for God's salvation by attending religious services at the local church every week. (Participant 3)</p>
	High levels of intrinsic religiousness help reduce depressive symptoms	<p>At that time, I felt that God was the ultimate meaning of my life. My faith in the belief that God loved me supported me in overcoming my physical ailments and emotional difficulties. Every time I prayed to God silently, I could feel the warmth and energy endowed by God in my heart, and those feelings indeed helped me a lot in my recovery from depression. (Participant 2)</p> <p>I believe that it was my unshakable faith in God that saved me from those years of depression." (Participant 4)</p>
Paradigms of treatment	The lack of effectiveness of medication	<p>It (taking drugs) was relatively effective in enabling me to fall asleep, but occasionally I still suffer from migraine when I work for a long time during the day. Also, I still sometimes come up with suicidal thoughts. I think medical treatments can indeed recover my body from the physical symptoms, but inside my heart, I'm still suffering sometimes. (Participant 3)</p>
	A need for spirituality-based therapeutic criteria	<p>My parents told me that my religion was against physiological intervention or medical treatment for diseases, so all I could do was meditate. (Participant 5)</p>

Discussion

Summary of Results

The primary purpose of this study was to examine intrinsic religiousness as a moderator of the correlations between religious practices, including organizational activities and non-organizational activities, and depression. In the study, the analyses of quantitative data indicated that the correlations between religious activities and depressive symptoms indeed depend on the levels of participants' intrinsic religiousness.

Consistent with past research, the results of this study demonstrated that high levels of intrinsic religiousness could strengthen the association between private religious or spiritual practices, such as prayer and meditation, and a reduction in depressive symptoms (Fehring, Miller, & Shaw, 1997).

However, unexpectedly, the regression analyses showed positive correlations between organizational religious activities and the levels of depression. This result was incompatible with the majority of previous research in religious practices and mental health, which suggested an inverse relationship between religiosity or spirituality and depressive symptoms (Smith, McCullough, & Poll, 2003). In the "Explaining the Results" section, several possible explanations were proposed to account for such a finding. Yet, the detailed mechanism underlying the positive correlations between organizational religious activities and depressive symptoms still remains unclear and unsure. Therefore, more sophisticated studies are required to analyze and explain the moderator models found in this study.

Limitations

Several limitations tempered the results of the present study. Firstly, this study was limited by its rather unrepresentative sample in both the quantitative phase and qualitative phase. The moderator model of religious practices, depression, and intrinsic religiousness established on the basis of the quantitative data collected could be gender-biased since all the participants recruited were female, and the relationship between the measured variables might differ in a sample of males. Secondly, the results might also be cultural-biased. This was because all the participants in the quantitative phase were recruited from the United States, a multicultural society consisting of a diversity of ethnicities and religions. Yet, the results might differ in the sample from a monocultural state, such as Japan or China. Another limitation was that the sample sizes of some ethnic groups were not large enough. For example, in the sample of the quantitative phase, there was only one American Indian or Alaska Native (0.5%) and four Black or African Americans (1.9%). Because of this, the ethnic minorities groups (except for the White or Caucasian group) were combined into one large group, the "Non-white" group. During the analysis of data, the overall sample was hence divided into two groups, the White and the Non-white group. This might be a problem because the model for each ethnic minority group could be different, yet combining them into one large group dismisses the distinctiveness across the religious traditions in different ethnic groups. Moreover, in the quantitative phase, the participants' religious beliefs were not measured, which meant that the difference in the moderator model across different religious traditions could not be investigated.

The same problems existed in the qualitative phase. Since all respondents were Asian or Asian American female Christians, the conclusion made based on the

qualitative data might lack representativeness and ecological validity, as the results could not be generalized to other genders, religions, or ethnic groups.

Implications

Given that the unrepresentative sample collected was a limitation of this study, future studies can examine a larger sample size consisting of more diverse genders, ethnicities, cultural backgrounds, and religious traditions. In addition, since the mechanisms underlying the model of the relationship between religious practices and depressive symptoms with intrinsic religiousness as the moderator remained unknown, future studies could focus on explaining such models.

Despite the limitations in the methodologies of this study, the results still demonstrated relevance to health psychology and the treatments of affective disorders. For instance, since the regression analyses indicated that non-organizational religious activities were associated with a significant reduction in depressive symptoms at high levels of intrinsic religiousness, the effectiveness of religiosity-based or spirituality-based therapeutic criteria for major depression disorders was confirmed. This type of psychotherapy might be especially useful for patients who were highly faithful in their religious beliefs yet whose religious doctrine forbids medication or surgical interventions as treatment for psychiatric disorders.

Moreover, the independent sample t-test conducted to compare the descriptive statistics of the White group and the Non-white group suggested a significant difference in educational level and depression level across the two groups. The Non-white group generally received a lower level of education compared to the White group and demonstrated a more significant tendency towards depressive symptoms. This finding might be insightful for public policymaking to promote equality in education and healthcare for ethnic minorities.

Concluding Comments

To conclude, the current researches related to religiousness were still not mature enough (Tix & Frazier, 2005). A number of the more interesting questions with regards to religiousness and mental health conditions have yet to be investigated. For example, what exact aspects of organizational or non-organizational practices are responsible for their influence on mental health? What psychological processes are involved in religious practices? How are the followers of distinct religious practices psychologically different from each other? Hopefully, a more in-depth understanding of this research domain may be acquired through additional, more sophisticated future research.

References

- AbdAleati, N. S., Mohd Zaharim, N., & Mydin, Y. O. (2014). Religiousness and mental health: Systematic review study. *Journal of Religion and Health, 55*(6), 1929-1937. doi:10.1007/s10943-014-9896-1
- Allen, G. E., & Wang, K. T. (2014). Examining religious commitment, perfectionism, scrupulosity, and well-being among individuals. *Psychology of Religion and Spirituality, 6*(3), 257-264. doi:10.1037/a0035197

- Balbuena, L., Baetz, M., & Bowen, R. (2013). Religious attendance, spirituality, and major depression in Canada: A 14-YEAR follow-up study. *The Canadian Journal of Psychiatry, 58*(4), 225-232. doi:10.1177/070674371305800408
- Barton, Y. A., Miller, L., Wickramaratne, P., Gameroff, M. J., & Weissman, M. M. (2013). Religious attendance and social adjustment as protective against depression: A 10-year prospective study. *Journal of Affective Disorders, 146*(1), 53-57. doi:10.1016/j.jad.2012.08.037
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck DEPRESSION inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*(1), 77-100. doi:10.1016/0272-7358(88)90050-5
- Desrosiers, A., & Miller, L. (2007). Relational spirituality and depression in adolescent girls. *Journal of Clinical Psychology, 63*(10), 1021-1037. doi:10.1002/jclp.20409
- Donahue, M. J. (1985). Intrinsic and extrinsic religiousness: Review and meta-analysis. *Journal of Personality and Social Psychology, 48*(2), 400-419. doi:10.1037/0022-3514.48.2.400
- Elkins, D. N., Hedstrom, L. J., Hughes, L. L., Leaf, J. A., & Saunders, C. (1988). Toward a humanistic-phenomenological spirituality. *Journal of Humanistic Psychology, 28*(4), 5-18. doi:10.1177/0022167888284002
- Emblen, J. D. (1992). Religion and spirituality defined according to current use in nursing literature. *Journal of Professional Nursing, 8*(1), 41-47. doi:10.1016/8755-7223(92)90116-g
- Esquivel, J. C. (2021). Religiously disaffiliated, religiously indifferent, or believers without religion? *Religions, 12*(7), 472. doi:10.3390/rel12070472
- Fehring, R. J., Miller, J. F., & Shaw, C. (1997). Spiritual well-being, religiosity, hope, depression, and other mood states in elderly people coping with cancer. *Oncol Nurs Forum, 24*, 663-671.
- Funder, D. C. (2002). Why study religion? *Psychological Inquiry, 13*, 213-214.
- Gorsuch, R. L. (2019). Toward motivational theories of intrinsic religious commitment. *The Psychology of Religion, 11*-24. doi:10.4324/9780429495915-2
- Gorsuch, R. L., & McPherson, S. E. (1989). Intrinsic/Extrinsic measurement: I/E-REVISED And SINGLE-ITEM Scales. *Journal for the Scientific Study of Religion, 28*(3), 348. doi:10.2307/1386745
- Graham, D. J. (2021). Religion data from Australia's 2016 census: The case of Judaism with broader implications for comparisons with earlier census results. *Journal for the Scientific Study of Religion, 60*(4), 789-810. doi:10.1111/jssr.12744
- Grosby, S. (2021). Religion and nationality in Europe. *The Oxford Handbook of Religion and Europe*, 155-174. doi:10.1093/oxfordhb/9780198834267.013.8
- Joiner, T. E., Perez, M., & Walker, R. L. (2002). Playing devil's advocate: Why not conclude that the relation of religiosity to mental health reduces to mundane mediators? *Psychological Inquiry, 13*, 214-216.
- King, M., Marston, L., McManus, S., Brugha, T., Meltzer, H., & Bebbington, P. (2013). Religion, spirituality and mental HEALTH: Results from a national study of English households. *British Journal of Psychiatry, 202*(1), 68-73. doi:10.1192/bjp.bp.112.112003
- Koenig, H. G. (1992). Religious coping and depression among elderly, hospitalized medically ill men. *American Journal of Psychiatry, 149*(12), 1693-1700. doi:10.1176/ajp.149.12.1693

- Koenig, H. G., & Büssing, A. (2010). The Duke University RELIGION INDEX (durel): A Five-Item measure for use IN Epidemiological studies. *Religions, 1*(1), 78-85. doi:10.3390/rel1010078
- Levin, J. S., & Markides, K. S. (1988). Religious attendance and psychological well-being in middle-aged and OLDER Mexican Americans. *Sociological Analysis, 49*(1), 66. doi:10.2307/3711104
- Loue, S. (2017). Religion, spirituality, mental health, and mental illness. *Handbook of Religion and Spirituality in Social Work Practice and Research, 143-161*. doi:10.1007/978-1-4939-7039-1_9
- McCullough, M. E., & Larson, D. B. (1999). Religion and depression: A review of the literature. *Twin Research, 2*(2), 126-136. doi:10.1375/twin.2.2.126
- McCullough, M. E., Hoyt, W. T., Larson, D. B., Koenig, H. G., & Thoresen, C. (2000). Religious involvement and mortality: A meta-analytic review. *Health Psychology, 19*(3), 211-222. doi:10.1037/0278-6133.19.3.211
- Miller, L., Wickramaratne, P., Gameroff, M. J., Sage, M., Tenke, C. E., & Weissman, M. M. (2012). Religiosity and major depression in adults at High risk: A Ten-Year prospective study. *American Journal of Psychiatry, 169*(1), 89-94. doi:10.1176/appi.ajp.2011.10121823
- Mosqueiro, B. P., Da Rocha, N. S., & Fleck, M. P. (2015). Intrinsic religiosity, resilience, quality of life, and suicide risk in depressed inpatients. *Journal of Affective Disorders, 179*, 128-133. doi: 10.1016/j.jad.2015.03.022
- Oman, D., & Lukoff, D. (2018). Mental health, religion, and spirituality. *Why Religion and Spirituality Matter for Public Health, 225-243*. doi:10.1007/978-3-319-73966-3_13
- Peterman, A. H., Fitchett, G., Brady, M. J., Hernandez, L., & Cella, D. (2002). Measuring spiritual well-being in people with cancer: The functional assessment of chronic ILLNESS therapy—spiritual Well-being scale (FACIT-Sp). *Annals of Behavioral Medicine, 24*(1), 49-58. doi:10.1207/s15324796abm2401_06
- Schnitzer, P. K., Brown, G. W., & Harris, T. (1980). Social origins of Depression: A study of psychiatric disorder in women. *Contemporary Sociology, 9*(1), 123. doi:10.2307/2065631
- Smith, T. B., McCullough, M. E., & Poll, J. (2003). Religiousness and depression: Evidence for a main effect and the moderating influence of stressful life events. *Psychological Bulletin, 129*(4), 614-636. doi:10.1037/0033-2909.129.4.614
- Tix, A. P., & Frazier, P. A. (2005). Mediation and moderation of the relationship between intrinsic religiousness and mental health. *Personality and Social Psychology Bulletin, 31*(3), 295-306. doi:10.1177/0146167204271592
- Tsai, J., Harpaz-Rotem, I., Pietrzak, R., & Southwick, S. (2012). The role of COPING, resilience, and social support in mediating the relation BETWEEN PTSD and social functioning in veterans returning from Iraq and Afghanistan. *PsycEXTRA Dataset*. doi:10.1037/e533652013-437
- VanderWeele, T. J., Jackson, J. W., & Li, S. (2016). Causal inference and longitudinal data: A case study of religion and mental health. *Social Psychiatry and Psychiatric Epidemiology, 51*(11), 1457-1466. doi:10.1007/s00127-016-1281-9
- Vaughan, F., Wittine, B., & Walsh, R. (1988). Transpersonal psychology and the religious person. *Religion and the Clinical Practice of Psychology., 483-509*. doi:10.1037/10199-018

- World Health Organization QOL Group. (1998). The World Health Organization quality of life ASSESSMENT (WHOQOL): Development and GENERAL psychometric properties. *Social Science & Medicine*, 46(12), 1569-1585. doi:10.1016/s0277-9536(98)00009-4
- Worthington, E. L., Kurusu, T. A., McCollough, M. E., & Sandage, S. J. (1996). Empirical research on religion and psychotherapeutic processes and outcomes: A 10-year review and research prospectus. *Psychological Bulletin*, 119(3), 448-487. doi:10.1037/0033-2909.119.3.448



Structure-Based Mixed Cannabinoid Design Presents New Opportunities for Fibrosis Treatment: A Computational Study

Sangwon Shin

Author Background: *Sangwon Shin grew up in Korea and currently attends SMIC Private School in Shanghai, China. His Pioneer research concentration was in the field of chemistry/biology and titled "Using AutoDock Vina to Design Next-Generation Small Molecule Therapeutics."*

Abstract

Fibrosis is a response to chronic injury in which parenchymal cells are replaced with extracellular matrix (ECM) components. Though associated with numerous diseases, there has yet been little clinical success for fibrosis treatment or reversal therapeutics. Cannabinoid receptor type 1 (CB1) and type 2 (CB2) have been established to have important regulatory roles in fibrosis. In particular, in-vitro and mouse model studies found CB1 antagonists and CB2 agonists effective in fibrosis attenuation. Therefore, a mixed CB1 antagonist/CB2 agonist may have significant therapeutic potential against fibrosis. Among such compounds, only URB447 is confirmed to bind to both CB1 and CB2 orthosteric sites with in-vivo effect. However, URB447 only exhibits weak binding affinities to both receptors. Hence, this study aimed to develop novel high-affinity synthetic cannabinoids with mixed CB1 antagonist/CB2 agonist properties using molecular docking, while examining the viability for in-silico structure-based synthetic cannabinoid design methods. Through analyzing interactions between URB447 analogs and CB1/CB2 residues, π - π stacking interactions were found to be a major factor in binding affinities of URB447 analogs. Furthermore, based on these interactions, two novel compounds, SWS001 and SWS002 were designed, a likely high-affinity mixed CB1 antagonist/CB2 agonist and a selective high-affinity CB2 agonist, respectively. In summary, this paper demonstrates that in-silico structure-based design is a viable, if not essential, step for the systematic development of mixed CB1/CB2 drugs.

1. Introduction

Fibrosis is a response to chronic injury in which parenchymal cells are replaced with extracellular matrix (ECM) components.¹ While it is an adaptive wound-healing and tissue repair process, it poses serious risks of intensifying into more lethal diseases.¹ Fibrotic diseases incur more than 800,000 mortalities per year,^{2,3} yet there has yet been little clinical success for fibrosis treatment or reversal therapeutics.^{1,3–5}

1.1. Etiology of Fibrosis

During the progression of fibrosis, fibroblasts—a common and differentiable cell lineage under normal conditions—excessively proliferate and differentiate into myofibroblasts, the predominant effector cell for fibrosis.⁶ Myofibroblasts, along with other activated cells, synthesize abnormally upregulated amounts of profibrotic markers, including smooth muscle α -actin 2 (ACTA2), transforming growth factor β (TGF- β), growth factors, matrix degradation inhibitors chemokines such as IL-1 β , IL-6, and IL-13, and ECM components such as collagen.^{6–8} These markers confer myofibroblasts with resistance to apoptosis and with time cause tissue scarring, further complicating the treatment of fibrosis.⁹

The specific pathobiology of fibrosis ultimately depends on the etiology. Not only does fibrosis occur in a wide variety of tissues,¹ the etiology in each tissue are diverse; in hepatic fibrosis alone, hepatitis B virus-related liver disease,¹⁰ hepatitis C virus-related liver disease,¹¹ alcoholic liver disease (ALD),¹² non-alcoholic fatty liver disease,¹² autoimmune hepatitis,^{13,14} and primary biliary cirrhosis can all evoke the development of fibrosis.¹³ Thus, current treatments for fibrosis mostly rely on treating the underlying condition behind fibrosis, but not fibrosis itself.^{4,15,16}

Even so, the unifying pathobiology in hepatic, pulmonary, renal, and cardiac fibrosis can be summarized in the following events:¹⁷ chronic inflammation as a result of the tissue-repair process, activation and proliferation of myofibroblasts, deposition of ECM components, and development of cirrhosis.^{1,3,18–20} Therefore, inhibiting key mechanisms that regulate these events could elucidate targets that could reveal novel antifibrotic therapies against late-stage fibrosis.

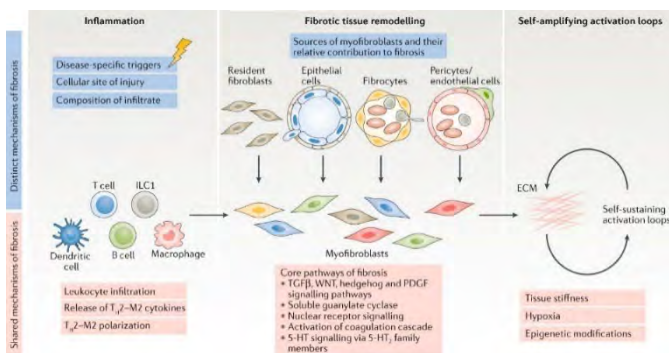


Figure 1. Derived from Distler, et al. Summary of fibrosis, with shared and distinct mechanisms between tissue types highlighted. Interestingly, as fibrosis progresses, symptoms and disease mechanisms from differing etiologies converge.

1.2. Reversing Fibrosis Through the Endocannabinoid System

To date, it is not completely known whether a reversal of fibrosis is possible.^{4,21} However, growing evidence in animal models and clinical studies suggests that fibrosis and subsequent cirrhosis are reversible after treatment of the underlying cause of chronic injury.^{4,21,22} One recently explored avenue is the targeting of the endocannabinoid system—a highly complex cell signaling pathway based on hydrophobic neurotransmitters called cannabinoids. This is primarily mediated by two class A G_{i/o}-coupled receptors (GPCR): cannabinoid receptor type 1 (CB1) and type 2 (CB2).^{8,23,24}

Numerous studies in the last two decades found that both CB1 mRNA and its receptor are significantly overexpressed in fibrotic tissues, particularly in myofibroblasts, highlighting its importance in fibrosis.^{22,25–27} Even so, the exact mechanism of the profibrotic effects of CB1 seems to be under continued scrutiny and differs between affected tissues,²⁸ though most studies agree that CB1 downregulation results in a more quiescent myofibroblast phenotype.^{7,8,29} A wide array of CB1 antagonists and inverse agonists are shown to inhibit the further progression of fibrosis both *in vitro* and *in vivo*.^{28–33}

Similarly, heightened CB2 expression is also correlated with the progression of fibrosis.^{7,24,27} In contrast to CB1, however, CB2 upregulation induced a decrease in profibrogenic chemokines, increased levels of collagenase, and modulated immune function—all principal hallmarks of fibrosis.²⁴ Overwhelming evidence through both *in vitro* and *in vivo* pharmacological studies using CB2 agonists suggests CB2 exerts an antifibrogenic effect.^{34–36} In summary, CB1 and CB2 regulate fibrosis in functionally opposite ways: CB1 by downregulation and CB2 by upregulation, both attenuate fibrosis.

1.3. Structural description of CB1 and CB2

1.3.1. Unifying features

First discovered in 1988,³⁷ CB1 is one of the most abundant receptors in the central nervous system (CNS).^{23,37–39} However, later studies revealed that CB1 also has a presence in the peripheral nervous system (PNS).^{38,40–43} Unlike CB1, CB2 is largely expressed within the immune system and hematopoietic cells.^{44,45} Though CB2 was originally assumed to be absent in the CNS at its discovery in 1993,⁴¹ later investigations found limited expression in the brain, particularly in the microglia.^{38,46–52}

So far, there have been six crystal and one cryo-EM structures of CB1, in complex with (1) antagonist (AM6538, 2.8Å),⁵³ (2) inverse agonist (taranabant, 2.6Å),⁵⁴ (3) agonist (CP55940, 3.25 Å),⁵⁵ (4) agonist (AM11542, 2.8 Å), (5) agonist (AM841, 2.9 Å), (6) allosteric modulator (ORG27569, 3.25 Å),³¹ and (7) agonist (MDMB-Fubinaca, 3Å).⁵⁶ On the other hand, there are fewer CB2 structural studies: an (1) antagonist (AM10257, 2.8Å)-bound CB2 crystal structure and a (2) agonist (WIN 55,212-2)-bound CB2 cryo-EM structure.^{57,58} Despite differences in biological function, CB2 has a high 44% amino acid sequence homology with CB1 with structural similarities.^{45,53,54,57,58}

Typical of GPCRs, CB1 and CB2 consists of 7 transmembrane α -helices (I-VII) and 1 amphipathic helix, with an extracellular N-terminus and an intercellular c-terminus.^{53–56,59} These helices are linked through three extracellular loops (ECLs) and 3 intracellular loops (ICLs) in both receptors.^{53–56,59} Furthermore, both receptors have a highly hydrophobic orthosteric pocket, containing the highly conserved disulfide bond within ECL2, which has been shown to play key roles in ligand-binding interactions.^{52–55,58} Overall, AM10257 and WIN 55,212-2 agonist-bound structures revealed that residues involved in CB2 agonist binding were nearly identical to those in CB1, potentially explaining the low receptor selectivity among cannabinoids.

1.3.2. CB1 and CB2 structural differences

Despite general similarities, important distinctions between CB1 and CB2 ligand interactions explain their different functional roles. In a comparison between CB2-AM10257 and CB1-AM6538 complexes' interactions, the binding of the CB2 agonist AM10257 is much more tightly constrained, while the CB1 antagonists AM6538 and taranabant are more extended.^{57,58} In addition, whereas AM10257 is bound to the receptor in a vertical conformation, AM6538 and taranabant are bound to the receptor more horizontally.^{57,58}

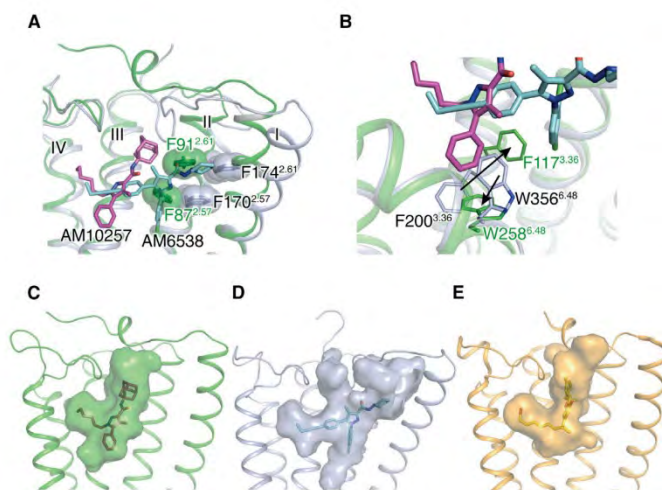


Figure 2. From Li, et al., (2019) [57](a) Superposition of the CB2-AM10257 and CB1-AM6538 ligand-binding pockets. The residues Phe2.57 and Phe2.61 are shown as spheres. AM10257-bound (magenta sticks) CB2 and AM6538-bound (cyan sticks) CB1 are shown as green and light blue cartoons, respectively. (b) Conformational difference of Phe3.36 and Trp6.48 in antagonist-bound CB2 and CB1. The color scheme is as in (A). (c-e) Surface representations of the ligand-binding pocket for CB2 with AM10257 (c), CB1 with AM6538 (d), and CB1 with AM11542 (e)

The crystal structure studies also highlighted the differences between CB1 and CB2 N termini. Unlike the CB1 N terminus, which inserts itself into the orthosteric site, the CB2 N terminus is significantly shorter and does not come into proximity with the orthosteric site or interact with the ligand.^{57,58} This difference is likely a key variable in understanding the different binding interactions of CB1 and CB2 ligands.

Finally, though both receptors retain the highly conserved Trp residue that functions as the “toggle switch,” its conformation in each receptor varies noticeably.^{57,58} The effects of these differences on specific ligands are unknown considering the highly variable nature of cannabinoid structures.

1.4. Principles of Synthetic Cannabinoid Classification

1.4.1. Classification

By definition, a cannabinoid is any compound that can bind and interact with any one of the cannabinoid receptors, including CB1 and CB2. Biologically, cannabinoids operate in two ways: cannabimimetic, compounds with physiological responses similar to those found in cannabis, and antagonists/inverse agonists that induce responses by blocking the receptors to other cannabinoids.⁶⁰

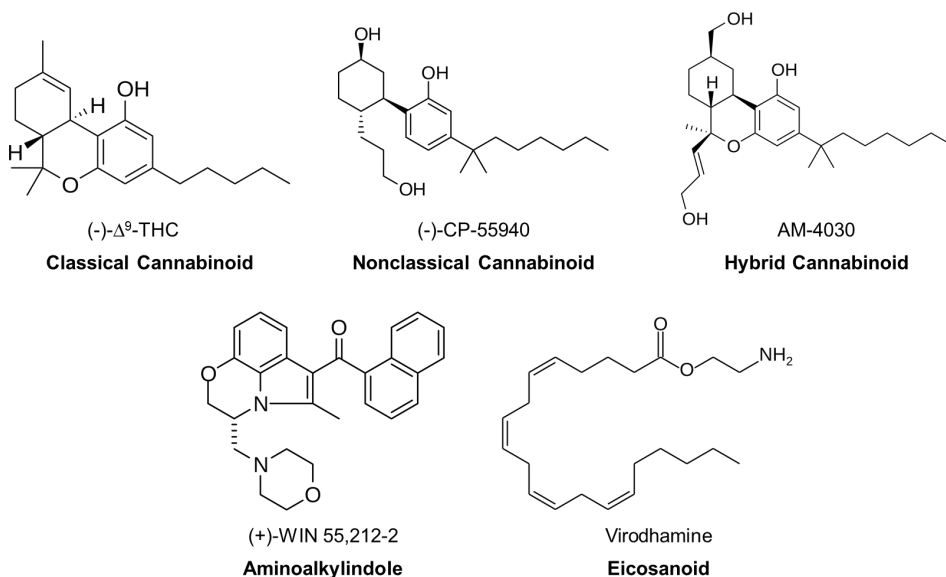


Figure 3. Classification of cannabinoids and schematic of example compounds

Since the discovery of trans- Δ^9 -tetrahydrocannabinol (THC), a highly diverse array of synthetic cannabinoids has been developed in the hopes of greater therapeutic functionality. Though this diversity makes classification of

cannabinoids difficult, five distinct classes have emerged over the past two decades:^{60,61}

1. *Classical cannabinoids* constitute THC; other cannabinoids from cannabis; and analogs from THC and other cannabinoids from cannabis.⁶⁰ Classical cannabinoids appear structurally similar to THC and are generally characterized by five structures: C3 side chain, phenolic hydroxyl, and three rings: the aromatic A-ring, southern aliphatic region (modifiable pyran B-ring), and cyclohexyl C-ring.⁶² These have three functional groups that interact with the receptor: (1) C11, (2) phenolic hydroxyl, and (3) side chain.⁶³
2. *Non-classical cannabinoids* are bicyclic cannabinoid analogs, 2-Cyclohexylphenol derivatives.^{60,64} In place of the modifiable pyran B-ring of classical cannabinoids, non-classical cannabinoids have a second aliphatic hydroxyl group; this group induces a heightened cannabimimetic activity.⁶⁴ Though structurally distinct from classical cannabinoids, non-classical cannabinoids failed to ameliorate the common problems of non-classical cannabinoids from the pharmaceutical perspective.^{65,66}
3. *Hybrid Cannabinoids* have the combined structural features of classical and non-classical cannabinoids.^{60,61,63,64} Typically, these compounds combine the B-ring of classical cannabinoids with the aliphatic hydroxyl group of non-classical cannabinoids, which forms a fourth functional group in addition to the three in classical cannabinoids: the southern aliphatic group (SAH).⁶³
4. *Aminoalkylindoles* is a highly diverse, strikingly different group of cannabinoids, originally characterized by the presence of an indole ring, with the nitrogen replaced by an aminoalkyl group.^{67,68} However, the former part of its name, aminoalkyl-, is no longer significant—the aminoalkyl group can be replaced by alkyl or aryl groups without affecting functionality.⁶⁰ The diversity and versatility of indole-based structures continue to give rise to novel synthetic cannabinoids, perhaps necessitating the development of a new classification system.⁶⁰
5. *Eicosanoids* contain the majority of endogenous cannabinoids (endocannabinoids), such as anandamide, and their synthetic analogs.⁶⁰ Structurally, eicosanoids are fatty-acid ethanolamides, far from typical THC-derived cannabinoids structures.^{69–71} Physiologically, they have a shorter onset and duration of action and affect cannabinoid-mediated pathways differently.⁶⁹ So far, however, eicosanoids are less understood than other cannabinoids, and only a few have been used therapeutically.⁷²
6. *Other Cannabinoids* include cannabinoids not suited to any other classification, such as pyrrole, indene, diarylpyrazole-based cannabinoids, naphthylmethylindenes or derivatives of naphthalen-1-yl-(4-pentyloxynaphthalen-1-yl)methanone.⁶⁰

1.4.2. Challenges to Effective Synthetic Cannabinoid Design

Despite the apparent structural simplicity of most cannabinoids, finding

synthetics both efficacious and devoid of undesirable side-effects has proven challenging.⁷³ This difficulty can likely be largely attributed to the wide range of signal transduction pathways of cannabinoid receptors and the structural homology between CB1 and CB2.

Ligands that exhibit functional selectivity induce differential receptor conformations to incur differing downstream effects.^{74,75} Given the expression of CB1 and CB2 in a wide variety of tissues, this concept is particularly relevant for cannabinoids.⁷⁶ Previous studies indicated that CB1 drugs can activate multiple conformations of CB1 that activate different Gi/Go subunits simultaneously with varying efficacy and potency.^{77–79} Drugs lacking functional selectivity, therefore, could potentially evoke undesirable side effects, such as sedation, tachycardia, panic, or depression-like symptoms.^{80–83} Fortunately, previous studies reported more favorable results for CB2 since there is surprisingly more functional selectivity for CB2 ligands.^{76,84–86} Nonetheless, the complexity of endocannabinoid receptor signaling presents a challenge to drug design: how can drugs that efficaciously and exclusively modulate one signaling pathway be designed?

Another difficulty in cannabinoid drug design lies within the homology of CB1 and CB2 receptors. Currently, a wide array of available CB1 and CB2 ligands can bind to both receptors without a strong bias for one receptor.⁸⁷ As an example, THC, now classified as a partial agonist, has a hCB1/hCB2 binding affinity ratio of only 0.7;⁸⁷ this is indicative of its failure as a clinical drug.⁸⁸ Ligands highly selective to one receptor have been found,^{80–83} such as APD371 (Olorinab), a full agonist to CB2, which has been successful in phase II clinical trials.⁹² Developing novel cannabinoids can be a challenging task, especially as *in vitro* successes do not always translate into clinical effectiveness.

1.4.3. Mixed CB1 and CB2 Synthetic Cannabinoids

Although it is advantageous to seek selectivity in drug design in most cases, drugs that bind multiple targets are often used situationally. For instance, given the opposing roles of CB1 and CB2, a drug that acts as a CB1 antagonist and a CB2 agonist may be useful.⁷³ For the CB1 and CB2 receptors, however, this is a relatively challenging task, likely due to their homology; so far, only a few studies seem to have serendipitously found synthetic cannabinoids that have such mixed target properties.

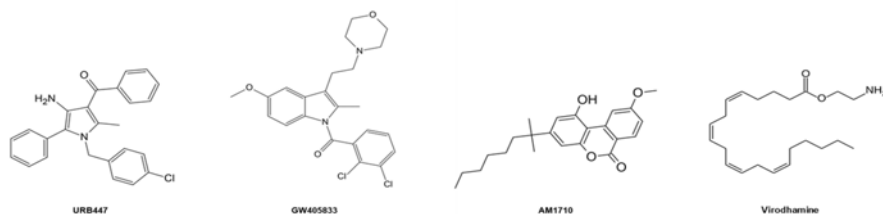


Figure 4. Two-dimensional structures of currently known mixed CB1 antagonist/CB2 agonist ligands

Only three synthetic and one endogenous mixed CB1 antagonist/CB2 agonists have been identified in relevant literature. The endogenous mixed CB1 antagonist/partial agonist and CB2 agonist is virodhamine,⁹³ which has been identified as a regulator of numerous other functional pathways.^{94–96} The first synthetic compound with such properties to be identified was URB447 ([4-amino-1-(4-chlorobenzyl)-2-methyl-5-phenyl-1H-pyrrol-3-yl](phenyl)methanone), a naphthoylpyrrole that acts as both a CB1 antagonist and CB2 agonist (see Figure 4).⁹⁷ In multiple studies, URB447 showed to have in vivo potency as an anti-obesity and neuroprotective drug without CNS activity, making it an ideal synthetic cannabinoid on which to base continued drug development.^{97,98}

A more recent study found two previously synthesized CB2 agonists that also showed efficacy in antagonizing CB1: GW405833 [(2,3-Dichlorophenyl) (5-methoxy-2-methyl-3-(2-morpholinoethyl)-1H-indol-1-yl)methanone] and AM1710 (1-hydroxy-9-methoxy-3-(2-methyloctan-2-yl)benzo[c]chromen-6-one) (see Figure 4).⁹⁹ GW405833 is a benzoylindole that was known to act as a partial agonist to CB2.^{99,100} Interestingly, current data suggest that GW405833 does not bind to the orthosteric site of CB1, though it is a high-potency, high-efficacy antagonist.⁹⁹ Given, however, that the allosteric site of GW405833 binding to CB1 has not yet been determined, it is difficult to use GW405833 as a model for the development of ligands with mixed properties.

AM1710 is a tricyclic ligand of the cannabidiol class, which has a carbonyl group in place of the 6,6-dimethyl moiety of the classical cannabinoid and an OCH₃ side chain (see Figure 4).⁹⁹ Unlike GW405833, AM1710 binds to the orthosteric site on CB1 as a low potency low-affinity competitive antagonist.⁹⁹ However, as URB447 is pyrrole-based and AM1710 is a classical cannabinoid, comparing ligand-receptor interactions between the two highly distinct structures is difficult. This makes elucidating the shared structural properties that allow URB447 and AM1710 to both antagonize CB1 and agonize CB2 a unique challenge. Furthermore, although AM1710 was highly potent in mouse and rat models, functionally relevant CB1 antagonism was not observed.¹⁰¹ Thus, though AM1710 is useful in rat and mouse model studies, it would not necessarily reflect human physiological responses.

Table 1. Binding affinities of synthetic mixed CB1 antagonist/CB2 agonist cannabinoids compared to THC, a partial agonist of both CB1 and CB2. Ratio of CB1/CB2 binding affinity was also displayed.

	Classification	CB1 (nM)	CB2 (nM)	CB1/CB2	Reference
THC	Classical	41 ± 2 ⁸⁸	36.4 ± 10 ⁸⁸	1.13	[88]
URB447	Naphthoylpyrrole	313 ± 72	41 ± 23	7.63	[78]
GW405833	Benzoylindole	4772 ± 1676	3.92 ± 1.58	1217	[99]
AM1710	Cannabidiol	360 ± 30	6.7 ± 1.3	53.7	[99]

1.5. Statement of Purpose

In this paper, I present a novel compound, SWS001, with both CB1 antagonist and CB2 agonist properties through in silico docking. Furthermore, this study fills the vacuum of mixed CB1 antagonist/CB2 agonist research and fulfills the need for a systematic approach for the design of such cannabinoids. Here, by

molecular docking and structure comparison, I raise and answer a crucial question that would improve mixed-cannabinoid design modalities: How can mixed-target cannabinoids be systematically designed in silico?

2. Materials and Methods

2.1. General Materials

This study was completed using the following software tools: Avogadro,¹⁰² UCSF ChimeraX,¹⁰³ AutoDock Vina,¹⁰⁴ ChemDraw, and Molinspiration Property Calculator.¹⁰⁵

2.1. Computational Analysis

2.1.1. Ligand Preparation

Rough geometry 3-dimensional structures of URB447 and its analogs were generated in Autodock Vina. Minimization of energy was done using the Autodock Vina auto-optimization tool, using the MMFF94 force field.¹⁰⁶ These energy-optimized ligands were saved as .pdb files to be read in ChimeraX.

2.1.2. Docking Preparation of CB1 and CB2 Receptor Proteins

CB1 and CB2 receptor proteins complexes with pre-docked ligands taranabant and AM10527, respectively, were retrieved from RCSB Protein DataBank (CB2: 5ZTY, CB1: 5U09), and imported to Chimera for dockprepping. Hydrogens were added and all nonstandard residues on the proteins were removed, including the pre-docked ligands, for convenience during docking. The ligand for docking was then imported from a .pdb file. The resulting dockprepped protein and the ligand were saved as a .py and .mol (separately for ligand and receptor) for later docking.

2.1.3. Docking

Analogs were docked using Autodock Vina within UCSF ChimeraX. The widely recognized orthosteric site was used as the binding cavity for both receptors. All default settings were used for docking. A grid box with a center at 20.3105, 2.13699, -7.7814 and a size of 25.069, 29.1168, 23.7918 was used for the CB1 receptor, and a grid box with a center at 7.63016, 2.62787, -57.5137 and a size of 25.5453, 23.5462, 24.0311 was used for the CB2 receptor.

To minimize possible human or software errors, docking was repeated a minimum of 10 times for each ligand. The best binding pose was determined by both docking score and visual analysis. The average docking score for the best binding pose was then calculated as the final docking score. ChimeraX tools were used to find hydrogen bonds and contacts/clashes and to create figures used in this study.

2.1.4. Calculation of Pharmacological Properties

Most pharmacological properties used in this study, LogS, LogP, and MW, were calculated using ChemDraw. TPSA was calculated through the Molinspiration Property Calculator.

3. Results

3.1. URB447 is a suitable model synthetic cannabinoid for analog development

In addition to being the only mixed CB1 antagonist/CB2 agonist that has demonstrated in vivo efficacy, several structural and chemical features make URB447 an excellent model for analog development.

First, molecular docking results demonstrated that URB447 binds to CB1 and CB2 in the same binding conformations as those described in crystal structure studies. Like taranabant and AM6538, URB447 is bound to CB1 in a horizontal, expanded conformation. On CB2, URB447 also mimicked the binding pose of AM10257, with a vertical but compact conformation. This similarity in binding pose to previously confirmed cannabinoid structures validated URB447 and its analogs as CB1 and CB2 orthosteric binders.

Second, URB447 is stable and structurally simple, notably based on the clinically used rimonabant. The number of Rotational Bonds (nRotB) of URB447 is only 5, making it resistant to significant torsion. Its three distinct arms that protrude from the relatively small pyrrole core are, in essence, phenyl groups with slight modifications. With the absence of large functional groups, URB447 provides the space needed for analog design within the tight orthosteric sites of CB1 and CB2. URB447 is also relatively small, with a large proportion of π -bonds, limiting the torsion often associated with the addition of functional groups. Hence, URB447 provides a “blank canvas” for analog design.

Third, URB447 structure is a soluble compound without CNS-mediated side effects. Calculated computationally, LogS, LogP, and TPSA values were -8.243, 5.52, and 46.33, respectively. This makes it near, if not within, the accepted values of a typical small molecule therapeutic. Combined, these factors support the rationale that URB447 is a logical model for analog development.

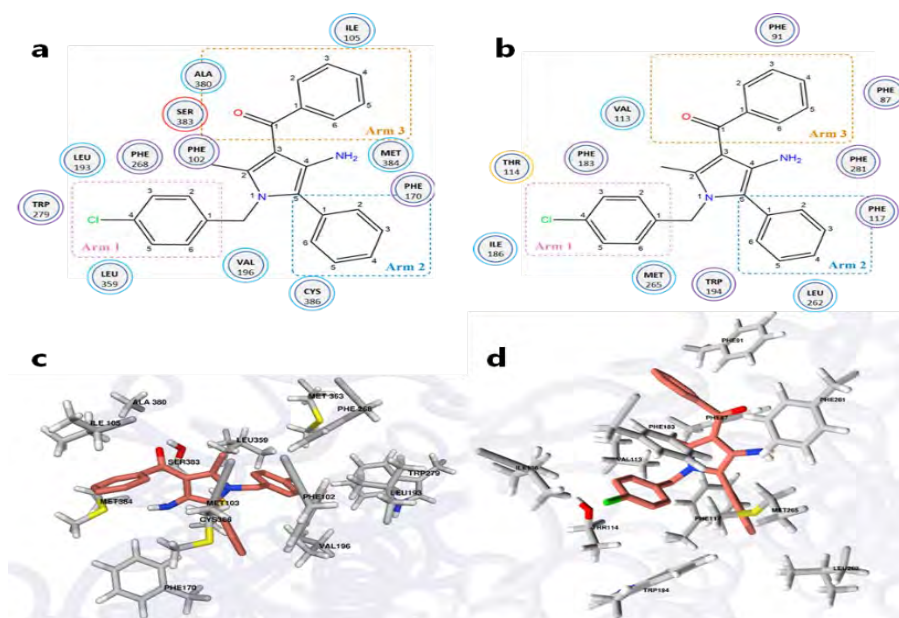


Figure 5. (a) two dimensional URB447 binding map with CB1. Blue rings represent Van Der Waal interactions, yellow rings represent polar interactions, red rings represent hydrogen bonds, and purple rings represent π - π interactions. (2) two dimensional URB447 binding map with CB2. (c) molecular docking of URB447 in CB1; residues within 4Å displayed. (d) molecular docking of URB447 in CB1; residues within 4Å displayed.

3.2. URB447 analogs were designed based on principles from CB1 and CB2 crystal structures and pre-existing drugs

Hence, URB447 was used as a model for ligand development in this study. Structural patterns of pyrrole-based or other ligands were further noted and tested on URB447 analogs in this study. Notably, unlike most synthetic cannabinoid characterization studies, this study heavily utilized the recent crystal structures of CB1 and CB2 for ligand-binding optimization. So far, ligands developed in such a way are uncommon, given that these crystal structure studies were published recently.

All ligands presented in this paper, therefore, have the same pyrrole core as URB447 and share numerous similar structural patterns, including the amino side chain and the propanone on arm 3. Before ligand modification, key ligand interactions for URB447 on both CB1 and CB2 were also identified and used as a basis for later docking analyses.

Through molecular docking, URB447 interactions with CB1 and CB2 residues were analyzed. This found that Van der Waals forces and π - π stacking were predominant interactions in CB1- and CB2-URB447 binding. These key π - π interactions were largely facilitated by the aromatic rings of URB447 and the numerous Phe and Trp residues (CB1: Phe170, Phe102, Phe268; CB2: Phe183,

Trp194, Phe117, Phe281, Phe87, Phe91). Comparatively, polar interactions were uncommon, and only one hydrogen bond was observed at residue Ser383 in CB1 with the propanone on arm 3. Given the importance of the π - π stacking in URB447, ligand design in this study focused on maximizing the effects of these interactions.

Moreover, to develop high-affinity analogs with clinical potential, four additional criteria were considered:

- (1) Is the drug's binding affinity higher than URB447 for both CB1 and CB2?
- (2) Is the ratio of the drug's binding affinity for CB1 and CB2 close to equivalent?
- (3) Is the drug soluble?
- (4) Is the drug peripherally restricted?

Criteria (1) was to ensure that ligands in this study improved binding affinity to both receptors, compared to URB447. Criteria (2) was due to the concern that URB447's binding affinity ratio (-10.6 to -11.2) between CB1 and CB2 was overly large. A drug with a higher overall binding affinity and a smaller ratio between the two receptors would be a more ideal mixed-target synthetic cannabinoid. Given the relatively hydrophobic nature of cannabinoids, Criteria (3) was essential, as drugs with higher binding affinity could be highly insoluble. Finally, as mentioned in the introduction, Criteria (4), peripheral restriction, is important for avoiding harmful CNS-mediated side-effects. To fulfill all four criteria, polar groups were added to lipophilic compounds with highest binding affinities, which were then reevaluated through docking.

3.3 π - π stacking interactions yielded high binding affinities in both CB1 and CB2

Based on the interactions observed from URB447 docking, modifications were introduced to have potent π - π stacking interactions. To achieve this, phenyl groups were added to two distinct positions of URB447.

The first modification of URB447 replaced the methyl group with a phenyl at C2 of the pyrrole core (R-group), resulting in compound **1**. In CB1, the phenyl group in place of the methyl group conferred additional interactions with aromatic and hydrophobic residues, such as Phe189, and Ile267, and likely established further π - π stacking with residues Phe102 and Phe279. The closer proximity to the CB1 ECL2 enabled π - π interactions with ECL2 residue Phe268. In CB2, potential π - π stacking was observed with Phe183, Phe91, Phe87, along with new interactions with residues Val113 and Ser90. Notably, the addition of the phenyl group caused a small, largely insignificant change in the binding pose of compound **1** compared to URB447 in both receptors. This relatively modest addition, however, improved the binding affinity against both receptors, from URB447's -10.6 kcal/mol to -12.0 kcal/mol in CB1 and -11.2 kcal/mol to -11.6 kcal/mol in CB2. Importantly, modifications without π bonds, such as trimethyl or adamantyl replacements failed to improve binding affinity.

The second modification of URB447 replaced the phenyl group on Arm 3 with a naphthyl group, resulting in compound **2**. The rationale for this

modification was in the abundance of residues able to form π - π stacking near C4 of Arm 3, such as His178 and Phe174 in CB1 and Phe106, Pro184, and Phe94 in CB2. Unlike in URB447 and compound **1**, these residues were well within 4Å of compound **2**. These additional interactions likely resulted in the high binding scores of -12.0 and -12.7 for CB1 and CB2, respectively. On the other hand, disruptions of these Van der Waals and π - π stacking interactions by adding large polar groups at the same arm yielded significantly lower binding affinities for both receptors. From compounds **2** and **3**, it is clear that π - π stacking interactions are correlated with high binding affinity scores in both CB1 and CB2.

3.4. The highly hydrophobic Compound 3 yielded the highest binding affinity

Given the successes of compounds **1** and **2** in moderately improving binding affinities for both receptors, features of both were combined to create compound **3**, which had both the methyl-to-phenyl group replacement in the R-group and the phenyl-to-naphthyl group replacement on Arm 3. In essence, therefore, Compound **3** initiated nearly all residue interactions that were observed in Compound **1** and **2**. As expected, compound **3** showed the strongest average binding affinities scores for both receptors among all compounds studied in this paper, with -12.5 and -12.1 for CB1 and CB2 respectively. It also exhibited a similar binding affinity to both CB1 and CB2, demonstrating potential as a mixed cannabinoid. The binding pose of the ligand also did not alter significantly from URB447.

However, the addition of two phenyl groups resulted in a highly hydrophobic compound. Compound **3**, therefore, was not a viable drug, due to its low solubility. In contrast to URB447, which had a logS of -8.243, compound **3** had a logS of -11.01. Considering other synthetic cannabinoids mostly have logS values of approximately -8, a compound with the structure of Compound **3** but with more polarity was desired. Hence, to improve binding affinity, polar elements were added to create a more soluble compound.

Interestingly, even though Compound **3** mostly exhibited a relatively equal binding affinity between the two receptors, most structures in Table S1 were selective to one receptor; this was especially surprising given that all these structures were analogs of compound **3**.

3.5. SWS001 is the strongest binding soluble mixed CB1 antagonist/CB2 agonist

SWS001 was the strongest binding compound that fully meets the four criteria listed above. The binding affinity for SWS001 for CB1 and CB2 were between URB447 and compound **3**, with -11.8 and -11.4 respectively. A logS value of -8.025 was also comparable to other clinical used cannabinoids, such as rimonabant. Furthermore, a high TPSA of 177.36 (in contrast to URB447's 48.03) indicates that it is likely peripherally restricted, thus minimizing the probability of off-target side-effects mediated by the CNS. Combined, these characteristics make SWS001 an ideal candidate for a mixed CB1

antagonist/CB2 agonist.

Compared to compound **3**, SWS001 holds three structural differences mainly introduced to increase solubility: that (1) certain carbons were replaced with nitrogen in strategic locations, (2) an oxygen atom was introduced in Arm 1, and (3) a chlorine atom was introduced in Arm 2. Docking results found that nitrogen atoms introduced into the phenyl and naphthyl rings reduced the binding affinity of SWS001 significantly compared to Compound **3**. After testing all possible combinations of carbon-to-nitrogen replacements within the Compound **3** structure, the optimal combination was found in SWS001. Predictably, the replacements in SWS001 were generally locations further from key binding interactions. In other words, the strategic replacement from carbon to nitrogen atoms in SWS001 maximized solubility while sacrificing the least amount of binding affinity. On the other hand, modification (2), did not yield significant decreases in binding affinity, despite a torsion in Arm 1. The introduced oxygen atom was instrumental in lowering the binding affinity to acceptable levels without a major loss in binding affinity. Modification (3) also maintained binding affinity and binding pose, while marginally increasing solubility. Hence, SWS001 retained the relatively high binding affinity of Compound **3** while favorably increasing solubility.

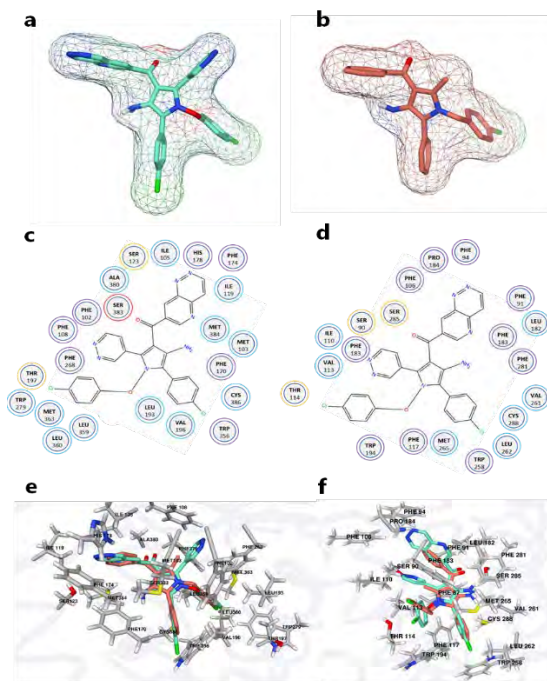


Figure 6. (a) Stick structure of SWS001 with solvent-excluded surface (SES) display. (b) Stick structure of URB447 with SES display. (c) two dimensional SWS001 binding map with CB1. See figure 1 caption for ring color representations. (d) two dimensional SWS001 binding map with CB2. (e) molecular docking of SWS001 with URB447 overlay in CB1; residues within 4Å displayed. (f) molecular docking of SWS001 with URB447 overlay in CB2; residues within 4Å displayed.

3.6. SWS002 is a likely high-affinity CB2 agonist

SWS002 was serendipitously discovered while creating URB447 analogs in search of mixed CB1 antagonist/CB2 agonist properties. In contrast to most ligands discussed in this paper, SWS002 exhibits a high degree of selectivity and affinity for CB2, with -8.8 kcal/mol in CB1 and -12.8 kcal/mol in CB2. Property simulation also found that SWS002 has a concerningly low logS value of -9.051 , indicating low solubility.

Structural differences between SWS002 and URB447 solely lie in Arm 3. However, unlike Compound 2, Arm 3 in SWS002 consists of a cyclohexane, an amide, and a pyridazine. Comparing with CB1, SWS002 has an almost identical binding conformation, with the cyclohexane aligning with the phenyl ring of URB447. From the cyclohexane, the pyridazine extends upwards, interacting with residues Ile105, Ile119, and Phe174 with π - π stacking and Van der Waals interactions. The amide extends downwards from the cyclohexane, forming a hydrogen bond with residue Ser123. Interestingly, it also forms hydrogen bonds with the protein backbone of Ile119. However, interactions with several other residues weaken SWS002-CB1 binding. The pyridazine ring interacts with residue His178, which in its charged conformation would repel

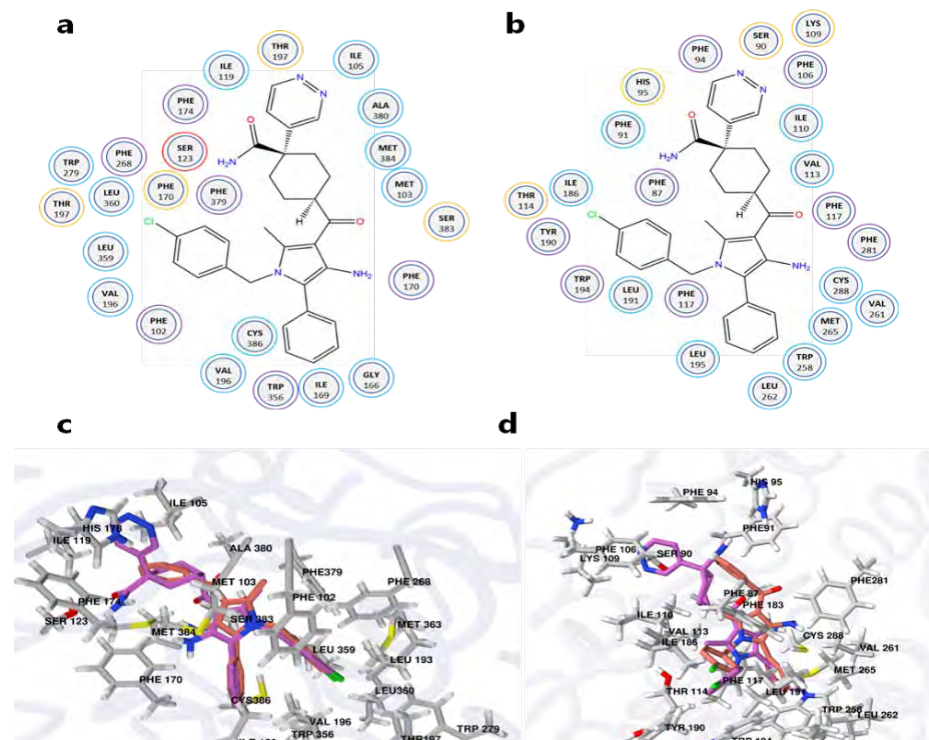


Figure 7. (c) two dimensional SWS002 binding map with CB1. See figure 1 caption for ring color representations. (d) two dimensional SWS002 binding map with CB2. (e) molecular docking of SWS002 with URB447 overlay in CB1; residues within 4Å displayed. (f) molecular docking of SWS002 with URB447 overlay in CB2; residues within 4Å displayed.

SWS002. Furthermore, the polar amide group possibly blocks essential π - π stacking and Van der Waals interactions with the pyridazine and cyclohexane rings. Eliminating nonpolar interactions with residues Met384, Phe174, and Phe170 may be a crucial factor in the weak binding affinity observed between SWS002 and CB1.

Unlike that seen in CB1, the pyridazine ring stretches horizontally, while the amide extends upwards. In a similar manner to SWS001, the pyridazine ring is positioned close to residues Phe106, Phe94, and Lys10 (though not Pro184), forming π - π stacking and polar interactions. The amide group also seems to form favorable interactions with residue His95, although hydrogen bonding was not observed. In addition, the shift in binding pose moved arm 1 forward, likely promoting greater π - π stacking interactions with Tyr190 and Trp194. Clashing was observed between SWS002 and residues Val113, Leu191, although binding affinity was unaffected.

4. Discussion

In this study, crystal structures of CB1 and CB2 were used to computationally design two novel synthetic cannabinoids—SWS001 and SWS002—with the goal of finding novel treatments for fibrosis.

SWS001 is a compound with a highly unique pharmacological property; it is a mixed cannabinoid, likely capable both as a high-affinity CB1 antagonist and a high-affinity CB2 agonist. So far, the only other cannabinoid with the same property and in vivo potency is URB447. However, SWS001 also resolves URB447's relatively low binding affinities to CB1 and CB2, while improving PSA and logS values. It is worth noting that the design process for SWS001 was both time-consuming and difficult; the hydrophobic nature of ligand-CB1 and -CB2 interactions required a delicate balance between maximizing binding affinity and solubility within the general structure of a small, relatively inflexible ligand. On the other hand, SWS002 is likely a highly selective, high-affinity CB2 agonist, though further modifications may be necessary to resolve its low solubility. Regardless, given that fibrosis is upregulated by CB1 and downregulated by CB2 activation, both compounds may have potential in fibrosis treatment.

Furthermore, as the only ligand design study primarily done using CB1 and CB2 crystal structures, this study confirmed and expanded on the key interactions necessary for cannabinoid binding. In particular, this study found that π - π stacking interactions between the ligand and Phe, His, and Pro residues greatly contribute to ligand binding. This is particularly true for URB447's arm 3 and its modifications, where nearby Phe, His, and Pro residues were especially abundant. Arm 3 also interacts with the N-terminus, a region predicted to be key to CB1's ligand binding affinity and the most noticeable difference between CB1 and CB2. Since arm 3 also seems to be the most variable part of the rimonabant-based structure, future rimonabant-based cannabinoid development should place a concerted effort in maximizing arm 3 π - π stacking interactions whilst maintaining enough solubility.

Even so, the computational and small-scale setting in this study creates an unmitigable limitation. Results from molecular docking simulations with an inactive conformation of CB1 and CB2 cannot be extrapolated to determine clinical potential. SWS001 and SWS002 should be further tested for functional selectivity, in vivo potency, CNS-mediated side effects, and aqueous solubility. Notably, though a mixed CB1 antagonist/CB2 agonist like SWS001 may have therapeutic properties against fibrosis, a lack of functional selectivity could cause broader unwanted effects. Hence, further validation in vitro and in vivo are required to ascertain SWS001 and SWS002's safety and effectiveness against fibrosis. Also, to confirm the importance of π - π stacking in CB1 and CB2 binding, particularly in arm 3, other analogs should be independently designed and tested. Similar levels of π - π stacking shown for an array of structurally diverse ligands would support such a conclusion.

Nonetheless, one definitive takeaway can be established from this study: despite the limitations, close analyses of protein structure and molecular docking are viable, if not essential, initial steps for mixed synthetic cannabinoid design.

References

- (1) Wynn, T. A. Cellular and Molecular Mechanisms of Fibrosis. *Journal of Pathology* **2008**, *214* (2), 199–210. <https://doi.org/10.1002/PATH.2277>.
- (2) Hinderer, S.; Schenke-Layland, K. Cardiac Fibrosis – A Short Review of Causes and Therapeutic Strategies. *Advanced Drug Delivery Reviews* **2019**, *146*, 77–82. <https://doi.org/10.1016/J.ADDR.2019.05.011>.
- (3) Murtha, L. A.; Schuliga, M. J.; Mabotuwana, N. S.; Hardy, S. A.; Waters, D. W.; Burgess, J. K.; Knight, D. A.; Boyle, A. J. The Processes and Mechanisms of Cardiac and Pulmonary Fibrosis. *Frontiers in Physiology* **2017**, *0* (OCT), 777. <https://doi.org/10.3389/FPHYS.2017.00777>.
- (4) Jung, Y. K.; Yim, H. J. Reversal of Liver Cirrhosis: Current Evidence and Expectations. *The Korean Journal of Internal Medicine* **2017**, *32* (2), 213–228. <https://doi.org/10.3904/KJIM.2016.268>.
- (5) Humphreys, B. D. Mechanisms of Renal Fibrosis. <https://doi.org/10.1146/annurev-physiol-022516-034227> **2018**, *80*, 309–326. <https://doi.org/10.1146/ANNUREV-PHYSIOL-022516-034227>.
- (6) Kendall, R. T.; Feghali-Bostwick, C. A. Fibroblasts in Fibrosis: Novel Roles and Mediators. *Frontiers in Pharmacology* **2014**, *0*, 123. <https://doi.org/10.3389/FPHAR.2014.00123>.
- (7) Lecru, L.; Desterke, C.; Grassin-Delyle, S.; Chatziantoniou, C.; Vandermeersch, S.; Devocelle, A.; Vernochet, A.; Ivanovski, N.; Ledent, C.; Ferlicot, S.; Dalia, M.; Saïd, M.; Beaudreuil, S.; Charpentier, B.; Vazquez, A.; Giron-Michel, J.; Azzarone, B.; Durrbach, A.; François, H. Cannabinoid Receptor 1 Is a Major Mediator of Renal Fibrosis. *Kidney International* **2015**, *88* (1), 72–84. <https://doi.org/10.1038/KI.2015.63>.

- (8) Teixeira-Clerc, F.; Julien, B.; Grenard, P.; van Nhieu, J. T.; Deveaux, V.; Li, L.; Serriere-Lanneau, V.; Ledent, C.; Mallat, A.; Lotersztajn, S. CB1 Cannabinoid Receptor Antagonism: A New Strategy for the Treatment of Liver Fibrosis. *Nature Medicine* 2006 12:6 **2006**, 12 (6), 671–676. <https://doi.org/10.1038/nm1421>.
- (9) Hinz, B.; Lagares, D. Evasion of Apoptosis by Myofibroblasts: A Hallmark of Fibrotic Diseases. *Nature Reviews Rheumatology* 2019 16:1 **2019**, 16 (1), 11–31. <https://doi.org/10.1038/s41584-019-0324-5>.
- (10) Mohamadnejad, M.; Montazeri, G.; Fazlollahi, A.; Zamani, F.; Nasiri, J.; Nobakht, H.; Forouzanfar, M. H.; Abedian, S.; Tavangar, S. M.; Mohamadkhani, A.; Ghoujehghi, F.; Estakhri, A.; Nouri, N.; Farzadi, Z.; Najjari, A.; Malekzadeh, R. Noninvasive Markers of Liver Fibrosis and Inflammation in Chronic Hepatitis B-Virus Related Liver Disease. *Official journal of the American College of Gastroenterology | ACG* **2006**, 101 (11).
- (11) Ghany, M. G.; Kleiner, D. E.; Alter, H.; Doo, E.; Khokar, F.; Promrat, K.; Herion, D.; Park, Y.; Liang, T. J.; Hoofnagle, J. H. Progression of Fibrosis in Chronic Hepatitis C. *Gastroenterology* **2003**, 124 (1), 97–104. <https://doi.org/10.1053/GAST.2003.50018>.
- (12) Zeybel, M.; Hardy, T.; Robinson, S. M.; Fox, C.; Anstee, Q. M.; Ness, T.; Masson, S.; Mathers, J. C.; French, J.; White, S.; Mann, J. Differential DNA Methylation of Genes Involved in Fibrosis Progression in Non-Alcoholic Fatty Liver Disease and Alcoholic Liver Disease. *Clinical Epigenetics* 2015 7:1 **2015**, 7 (1), 1–11. <https://doi.org/10.1186/S13148-015-0056-6>.
- (13) Chazouillères, O.; Wendum, D.; Serfaty, L.; Rosmorduc, O.; Poupon, R. Long Term Outcome and Response to Therapy of Primary Biliary Cirrhosis—Autoimmune Hepatitis Overlap Syndrome. *Journal of Hepatology* **2006**, 44 (2), 400–406. <https://doi.org/10.1016/J.JHEP.2005.10.017>.
- (14) Björnsson, E.; Talwalkar, J.; Treeprasertsuk, S.; Kamath, P. S.; Takahashi, N.; Sanderson, S.; Neuhauser, M.; Lindor, K. Drug-Induced Autoimmune Hepatitis: Clinical Characteristics and Prognosis. *Hepatology* **2010**, 51 (6), 2040–2048. <https://doi.org/10.1002/HEP.23588>.
- (15) Boor, P.; Ostendorf, T.; Floege, J. Renal Fibrosis: Novel Insights into Mechanisms and Therapeutic Targets. *Nature Reviews Nephrology* 2010 6:11 **2010**, 6 (11), 643–656. <https://doi.org/10.1038/nrneph.2010.120>.
- (16) Smith, M. L. Update on Pulmonary Fibrosis: Not All Fibrosis Is Created Equally. *Archives of Pathology & Laboratory Medicine* **2016**, 140 (3), 221–229. <https://doi.org/10.5858/ARPA.2015-0288-SA>.
- (17) Distler, J. H. W.; Györfi, A. H.; Ramanujam, M.; Whitfield, M. L.; Königshoff, M.; Lafyatis, R. Shared and Distinct Mechanisms of Fibrosis. *Nature Reviews Rheumatology* **2019**, 15 (12), 705–730. <https://doi.org/10.1038/S41584-019-0322-7>.
- (18) Liu, Y. Cellular and Molecular Mechanisms of Renal Fibrosis. *Nature Reviews Nephrology* 2011 7:12 **2011**, 7 (12), 684–696. <https://doi.org/10.1038/nrneph.2011.149>.
- (19) Roehlen, N.; Crouchet, E.; Baumert, T. F. Liver Fibrosis: Mechanistic

- Concepts and Therapeutic Perspectives. *Cells* 2020, Vol. 9, Page 875 **2020**, 9 (4), 875. <https://doi.org/10.3390/CELLS9040875>.
- (20) Thannickal, V. J.; Toews, G. B.; White, E. S. ; III, J. P. L.; Martinez, F. J. Mechanisms of Pulmonary Fibrosis. <https://doi.org/10.1146/annurev.med.55.091902.103810> **2004**, 55, 395–417. <https://doi.org/10.1146/ANNUREV.MED.55.091902.103810>.
- (21) Friedman, S. L.; Bansal, M. B. Reversal of Hepatic Fibrosis — Fact or Fantasy? *Hepatology* **2006**, 43 (S1), S82–S88. <https://doi.org/https://doi.org/10.1002/hep.20974>.
- (22) Kisseleva, T.; Brenner, D. A. Hepatic Stellate Cells and the Reversal of Fibrosis. *Journal of Gastroenterology and Hepatology* **2006**, 21 (SUPPL. 3), S84–S87. <https://doi.org/10.1111/J.1440-1746.2006.04584.X>.
- (23) Howlett, A. C.; Breivogel, C. S.; Childers, S. R.; Deadwyler, S. A.; Hampson, R. E.; Porrino, L. J. Cannabinoid Physiology and Pharmacology: 30 Years of Progress. *Neuropharmacology* **2004**, 47 (SUPPL. 1), 345–358. <https://doi.org/10.1016/J.NEUROPHARM.2004.07.030>.
- (24) Julien, B.; Grenard, P.; Teixeira-Clerc, F.; Nhieu, J. T. van; Li, L.; Karsak, M.; Zimmer, A.; Mallat, A.; Lotersztajn, S. Antifibrogenic Role of the Cannabinoid Receptor CB2 in the Liver. *Gastroenterology* **2005**, 128 (3), 742–755. <https://doi.org/10.1053/j.gastro.2004.12.050>.
- (25) Lecru, L.; Desterke, C.; Grassin-Delyle, S.; Chatziantoniou, C.; Vandermeersch, S.; Devocelle, A.; Vernochet, A.; Ivanovski, N.; Ledent, C.; Ferlicot, S.; Dalia, M.; Saïd, M.; Beaudreuil, S.; Charpentier, B.; Vazquez, A.; Giron-Michel, J.; Azzarone, B.; Durrbach, A.; François, H. Cannabinoid Receptor 1 Is a Major Mediator of Renal Fibrosis. *Kidney International* **2015**, 88 (1), 72–84. <https://doi.org/10.1038/KI.2015.63>.
- (26) H, F.; C, C. Renal Fibrosis: Recent Translational Aspects. *Matrix biology : journal of the International Society for Matrix Biology* **2018**, 68–69, 318–332. <https://doi.org/10.1016/J.MATBIO.2017.12.013>.
- (27) Dai, E.; Zhang, L.; Ye, L.; Wan, S.; Feng, L.; Qi, Q.; Yao, F.; Li, Z. Hepatic Expression of Cannabinoid Receptors CB1 and CB2 Correlate with Fibrogenesis in Patients with Chronic Hepatitis B. *International Journal of Infectious Diseases* **2017**, 59, 124–130. <https://doi.org/10.1016/J.IJID.2017.03.008>.
- (28) Cinar, R.; Gochuico, B. R.; Iyer, M. R.; Jourdan, T.; Yokoyama, T.; Park, J. K.; Coffey, N. J.; Pri-Chen, H.; Szanda, G.; Liu, Z.; Mackie, K.; Gahl, W. A.; Kunos, G. Cannabinoid CB1 Receptor Overactivity Contributes to the Pathogenesis of Idiopathic Pulmonary Fibrosis. *JCI Insight* **2017**, 2 (8). <https://doi.org/10.1172/JCI.INSIGHT.92281>.
- (29) Tan, S.; Liu, H.; Ke, B.; Jiang, J.; Wu, B. The Peripheral CB1 Receptor Antagonist JD5037 Attenuates Liver Fibrosis via a CB1 Receptor/ β -Arrestin1/Akt Pathway. *British Journal of Pharmacology* **2020**, 177 (12), 2830–2847. <https://doi.org/10.1111/BPH.15010>.
- (30) Marquart, S.; Zerr, P.; Akhmetshina, A.; Palumbo, K.; Reich, N.; Tomcik, M.; Horn, A.; Dees, C.; Engel, M.; Zwerina, J.; Distler, O.; Schett, G.; Distler, J. H. W. Inactivation of the Cannabinoid Receptor CB1 Prevents Leukocyte Infiltration and Experimental Fibrosis. *Arthritis &*

- Rheumatism* **2010**, *62* (11), 3467–3476. <https://doi.org/10.1002/ART.27642>.
- (31) Palumbo-Zerr, K.; Horn, A.; Distler, A.; Zerr, P.; Dees, C.; Beyer, C.; Selvi, E.; Cravatt, B. F.; Distler, O.; Schett, G.; Distler, J. H. W. Inactivation of Fatty Acid Amide Hydrolase Exacerbates Experimental Fibrosis by Enhanced Endocannabinoid-Mediated Activation of CB1. *Annals of the Rheumatic Diseases* **2012**, *71* (12), 2051–2054. <https://doi.org/10.1136/ANNRHEUMDIS-2012-201823>.
- (32) Lin, C. Y.; Hsu, Y. J.; Hsu, S. C.; Chen, Y.; Lee, H. S.; Lin, S. H.; Huang, S. M.; Tsai, C. S.; Shih, C. C. CB1 Cannabinoid Receptor Antagonist Attenuates Left Ventricular Hypertrophy and Akt-Mediated Cardiac Fibrosis in Experimental Uremia. *Journal of Molecular and Cellular Cardiology* **2015**, *85*, 249–261. <https://doi.org/10.1016/J.YJMCC.2015.06.010>.
- (33) Giannone, F. A.; Baldassarre, M.; Domenicali, M.; Zaccherini, G.; Trevisani, F.; Bernardi, M.; Caraceni, P. Reversal of Liver Fibrosis by the Antagonism of Endocannabinoid CB1 Receptor in a Rat Model of CCl4-Induced Advanced Cirrhosis. *Laboratory Investigation* **2012** *92*:3 **2011**, *92* (3), 384–395. <https://doi.org/10.1038/labinvest.2011.191>.
- (34) Akhmetshina, A.; Dees, C.; Busch, N.; Beer, J.; Sarter, K.; Zwerina, J.; Zimmer, A.; Distler, O.; Schett, G.; Distler, J. H. W. The Cannabinoid Receptor CB2 Exerts Antifibrotic Effects in Experimental Dermal Fibrosis. *Arthritis & Rheumatism* **2009**, *60* (4), 1129–1136. <https://doi.org/10.1002/ART.24395>.
- (35) Muñoz-Luque, J.; Ros, J.; Fernández-Varo, G.; Tugues, S.; Morales-Ruiz, M.; Alvarez, C. E.; Friedman, S. L.; Arroyo, V.; Jiménez, W. Regression of Fibrosis after Chronic Stimulation of Cannabinoid CB2 Receptor in Cirrhotic Rats. *Journal of Pharmacology and Experimental Therapeutics* **2008**, *324* (2), 475–483. <https://doi.org/10.1124/JPET.107.131896>.
- (36) Wawryk-Gawda, E.; Chłapek, K.; Zarobkiewicz, M. K.; Lis-Sochocka, M.; Chylińska-Wrzos, P.; Boguszewska-Czubara, A.; Sławiński, M. A.; Franczak, A.; Jodłowska-Jędrych, B.; Biała, G. CB2R Agonist Prevents Nicotine Induced Lung Fibrosis. <https://doi.org/10.1080/01902148.2018.1543368> **2019**, *44* (7), 344–351. <https://doi.org/10.1080/01902148.2018.1543368>.
- (37) Devane, W. A.; Dysarz, F. A.; Johnson, M. R.; Melvin, L. S.; Howlett, A. C. Determination and Characterization of a Cannabinoid Receptor in Rat Brain. *Molecular Pharmacology* **1988**, *34* (5), 605.
- (38) Herkenham, M.; Lynn, A. B.; Little, M. D.; Johnson, M. R.; Melvin, L. S.; de Costa, B. R.; Rice, K. C. Cannabinoid Receptor Localization in Brain. *Proceedings of the National Academy of Sciences* **1990**, *87* (5).
- (39) Matsuda, L. A.; Lolait, S. J.; Brownstein, M. J.; Young, A. C.; Bonner, T. I. *Structure of a Cannabinoid Receptor and Functional Expression of the Cloned EDNA*; 1990.
- (40) MC, S.-P.; NM, S.; K, M.; JM, W.; K, T. CB1 Receptor Localization in Rat Spinal Cord and Roots, Dorsal Root Ganglion, and Peripheral Nerve. *Zhongguo yao li xue bao = Acta Pharmacologica Sinica* **1999**, *20* (12), 1115–1120.
- (41) Gómez, R.; Navarro, M.; Ferrer, B.; Trigo, J. M.; Bilbao, A.; Arco, I. del;

- Cippitelli, A.; Nava, F.; Piomelli, D.; Fonseca, F. R. de. A Peripheral Mechanism for CB1 Cannabinoid Receptor-Dependent Modulation of Feeding. *Journal of Neuroscience* **2002**, *22* (21), 9612–9617. <https://doi.org/10.1523/JNEUROSCI.22-21-09612.2002>.
- (42) Nogueiras, R.; Veyrat-Durebex, C.; Suchanek, P. M.; Klein, M.; Tschöp, J.; Caldwell, C.; Woods, S. C.; Wittmann, G.; Watanabe, M.; Liposits, Z.; Fekete, C.; Reizes, O.; Rohner-Jeanrenaud, F.; Tschöp, M. H. Peripheral, but Not Central, CB1 Antagonism Provides Food Intake–Independent Metabolic Benefits in Diet-Induced Obese Rats. *Diabetes* **2008**, *57* (11), 2977–2991. <https://doi.org/10.2337/DB08-0161>.
- (43) Richardson, J. D.; Kilo, S.; Hargreaves, K. M. Cannabinoids Reduce Hyperalgesia and Inflammation via Interaction with Peripheral CB1 Receptors. *Pain* **1998**, *75* (1), 111–119. [https://doi.org/10.1016/S0304-3959\(97\)00213-3](https://doi.org/10.1016/S0304-3959(97)00213-3).
- (44) Griffin, G.; Fernando, S. R.; Ross, R. A.; McKay, N. G.; Ashford, M. L. J.; Shire, D.; Huffman, J. W.; Yu, S.; Lainton, J. A. H.; Pertwee, R. G. Evidence for the Presence of CB2-like Cannabinoid Receptors on Peripheral Nerve Terminals. *European Journal of Pharmacology* **1997**, *339* (1), 53–61. [https://doi.org/10.1016/S0014-2999\(97\)01336-8](https://doi.org/10.1016/S0014-2999(97)01336-8).
- (45) Munro, S.; Thomas, K. L.; Abu-Shaar, M. Molecular Characterization of a Peripheral Receptor for Cannabinoids. *Nature* **1993**, *365* (6441), 61–65. <https://doi.org/10.1038/365061a0>.
- (46) ONAIVI, E. S.; ISHIGURO, H.; GONG, J.-P.; PATEL, S.; PERCHUK, A.; MEOZZI, P. A.; MYERS, L.; MORA, Z.; TAGLIAFERRO, P.; GARDNER, E.; BRUSCO, A.; AKINSHOLA, B. E.; LIU, Q.-R.; HOPE, B.; IWASAKI, S.; ARINAMI, T.; TEASENFITZ, L.; UHL, G. R. Discovery of the Presence and Functional Expression of Cannabinoid CB2 Receptors in Brain. *Annals of the New York Academy of Sciences* **2006**, *1074* (1), 514–536. <https://doi.org/10.1196/ANNALS.1369.052>.
- (47) Callén, L.; Moreno, E.; Barroso-Chinea, P.; Moreno-Delgado, D.; Cortés, A.; Mallol, J.; Casadó, V.; Lanciego, J. L.; Franco, R.; Lluís, C.; Canela, E. I.; McCormick, P. J. Cannabinoid Receptors CB1 and CB2 Form Functional Heteromers in Brain *. *Journal of Biological Chemistry* **2012**, *287* (25), 20851–20865. <https://doi.org/10.1074/JBC.M111.335273>.
- (48) A. B.; PA, T.; T, S.; ES, O. Ultrastructural Localization of Neuronal Brain CB2 Cannabinoid Receptors. *Annals of the New York Academy of Sciences* **2008**, *1139*, 450–457. <https://doi.org/10.1196/ANNALS.1432.037>.
- (49) Gong, J. P.; Onaivi, E. S.; Ishiguro, H.; Liu, Q. R.; Tagliaferro, P. A.; Brusco, A.; Uhl, G. R. Cannabinoid CB2 Receptors: Immunohistochemical Localization in Rat Brain. *Brain Research* **2006**, *1071* (1), 10–23. <https://doi.org/10.1016/J.BRAINRES.2005.11.035>.
- (50) C, B.; WK, K.; I, C.; CJ, H.; K, M.; RM, T.; K, W.; J, R. A Glial Endogenous Cannabinoid System Is Upregulated in the Brains of Macaques with Simian Immunodeficiency Virus-Induced Encephalitis. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience* **2005**, *25* (10), 2530–2536. <https://doi.org/10.1523/JNEUROSCI.3923-04.2005>.

- (51) Onaivi, E. S. Neuropsychobiological Evidence for the Functional Presence and Expression of Cannabinoid CB2 Receptors in the Brain. *Neuropsychobiology* **2006**, *54* (4), 231–246. <https://doi.org/10.1159/000100778>.
- (52) FS, den B.; P, C.; Q, S.-Z.; W, van A.; M, B.; S, O.; CG, K.; M, M.; WJ, W.; TR, W. Excitability of Prefrontal Cortical Pyramidal Neurons Is Modulated by Activation of Intracellular Type-2 Cannabinoid Receptors. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109* (9), 3534–3539. <https://doi.org/10.1073/PNAS.1118167109>.
- (53) Hua, T.; Vemuri, K.; Pu, M.; Qu, L.; Han, G. W.; Wu, Y.; Zhao, S.; Shui, W.; Li, S.; Korde, A.; Laprairie, R. B.; Stahl, E. L.; Ho, J.-H.; Zvonok, N.; Zhou, H.; Kufareva, I.; Wu, B.; Zhao, Q.; Hanson, M. A.; Bohn, L. M.; Makriyannis, A.; Stevens, R. C.; Liu, Z.-J. Crystal Structure of the Human Cannabinoid Receptor CB1. *Cell* **2016**, *167* (3), 750–762.e14. <https://doi.org/10.1016/J.CELL.2016.10.004>.
- (54) Shao, Z.; Yin, J.; Chapman, K.; Grzemska, M.; Clark, L.; Wang, J.; Rosenbaum, D. M. High-Resolution Crystal Structure of the Human CB1 Cannabinoid Receptor. *Nature* **2016** *540*:7634 **2016**, *540* (7634), 602–606. <https://doi.org/10.1038/nature20613>.
- (55) Shao, Z.; Yan, W.; Chapman, K.; Ramesh, K.; Ferrell, A. J.; Yin, J.; Wang, X.; Xu, Q.; Rosenbaum, D. M. Structure of an Allosteric Modulator Bound to the CB1 Cannabinoid Receptor. *Nature Chemical Biology* **2019** *15*:12 **2019**, *15* (12), 1199–1205. <https://doi.org/10.1038/s41589-019-0387-2>.
- (56) Krishna Kumar, K.; Shalev-Benami, M.; Robertson, M. J.; Hu, H.; Banister, S. D.; Hollingsworth, S. A.; Latorraca, N. R.; Kato, H. E.; Hilger, D.; Maeda, S.; Weis, W. I.; Farrens, D. L.; Dror, R. O.; Malhotra, S. v; Kobilka, B. K.; Skiniotis, G. Structure of a Signaling Cannabinoid Receptor 1-G Protein Complex. *Cell* **2019**, *176* (3), 448–458.e12. <https://doi.org/10.1016/j.cell.2018.11.040>.
- (57) X, L.; T, H.; K, V.; JH, H.; Y, W.; L, W.; P, P.; O, B.; N, Z.; K, L.; L, Q.; GW, H.; MR, I.; R, C.; NJ, C.; J, W.; M, W.; V, K.; S, Z.; G, K.; LM, B.; A, M.; RC, S.; ZJ, L. Crystal Structure of the Human Cannabinoid Receptor CB2. *Cell* **2019**, *176* (3), 459–467.e13. <https://doi.org/10.1016/J.CELL.2018.12.011>.
- (58) Xing, C.; Zhuang, Y.; Xu, T. H.; Feng, Z.; Zhou, X. E.; Chen, M.; Wang, L.; Meng, X.; Xue, Y.; Wang, J.; Liu, H.; McGuire, T. F.; Zhao, G.; Melcher, K.; Zhang, C.; Xu, H. E.; Xie, X. Q. Cryo-EM Structure of the Human Cannabinoid Receptor CB2-Gi Signaling Complex. *Cell* **2020**, *180* (4), 645–654.e13. <https://doi.org/10.1016/J.CELL.2020.01.007>.
- (59) Bramblett, R. D.; Panu, A. M.; Ballesteros, J. A.; Reggio, P. H. Construction of a 3D Model of the Cannabinoid Cb1 Receptor: Determination of Helix Ends and Helix Orientation. *Life Sciences* **1995**, *56* (23–24), 1971–1982. [https://doi.org/10.1016/0024-3205\(95\)00178-9](https://doi.org/10.1016/0024-3205(95)00178-9).
- (60) Shevyrin, V. A.; Morzherin, Y. Y. Cannabinoids: Structures, Effects, and Classification. *Izvestiya Akademii Nauk. Seriya Khimicheskaya* **2015**, *64* (6), 1249–1266.
- (61) Howlett, A. C.; Abood, M. E. CHAPTER FIVE CB 1 and CB 2 Receptor Pharmacology. **2017**. <https://doi.org/10.1016/bs.apha.2017.03.007>.

- (62) Bow, E. W.; Rimoldi, J. M. The Structure–Function Relationships of Classical Cannabinoids: CB1/CB2 Modulation: <https://doi.org/10.4137/PMC.S32171> **2016**, *8*, 17–39. <https://doi.org/10.4137/PMC.S32171>.
- (64) Drake, D. J.; Jensen, R. S.; Busch-Petersen, J.; Kawakami, J. K.; Concepcion Fernandez-Garcia, M.; Fan, P.; Makriyannis, A.; Tius, M. A. Classical/Nonclassical Hybrid Cannabinoids: Southern Aliphatic Chain-Functionalized C-6 Methyl, Ethyl, and Propyl Analogues. **1998**.
- (65) Weissman, A.; Milne, G. M.; Melvin, L. S. Cannabimimetic Activity from CP-47,497, a Derivative of 3-Phenylcyclohexanol. *Journal of Pharmacology and Experimental Therapeutics* **1982**, *223* (2).
- (63) BR, M.; DR, C.; WR, P.; RL, B.; RK, R. Pharmacological Evaluation of Dimethylheptyl Analogs of Delta 9-THC: Reassessment of the Putative Three-Point Cannabinoid-Receptor Interaction. *Drug and Alcohol Dependence* **1995**, *37* (3), 231–240. [https://doi.org/10.1016/0376-8716\(94\)01081-U](https://doi.org/10.1016/0376-8716(94)01081-U).
- (66) Drake, D. J.; Jensen, R. S.; Busch-Petersen, J.; Kawakami, J. K.; Concepcion Fernandez-Garcia, M.; Fan, P.; Makriyannis, A.; Tius, M. A. Classical/Nonclassical Hybrid Cannabinoids: Southern Aliphatic Chain-Functionalized C-6 Methyl, Ethyl, and Propyl Analogues. **1998**.
- (67) Jansen, E. M.; Haycock, D. A.; Ward, S. J.; Seybold, V. S. Distribution of Cannabinoid Receptors in Rat Brain Determined with Aminoalkylindoles. *Brain Research* **1992**, *575* (1), 93–102. [https://doi.org/10.1016/0006-8993\(92\)90428-C](https://doi.org/10.1016/0006-8993(92)90428-C).
- (68) Bell, M. R.; D'Ambra, T. E.; Kumar, V.; Eissenstat, M. A.; Jr., J. L. H.; Wetzel, J. R.; Rosi, D.; Pillion, R. E.; Daum, S. J. Antinociceptive (Aminoalkyl)Indoles. *Journal of Medicinal Chemistry* **2002**, *34* (3), 1099–1110. <https://doi.org/10.1021/JM00107A034>.
- (69) Lin, S.; Khanolkar, A. D.; Fan, P.; Goutopoulos, A.; Qin, C.; Papahadjis, D.; Makriyannis, A. Novel Analogues of Arachidonylethanolamide (Anandamide): Affinities for the CB1 and CB2 Cannabinoid Receptors and Metabolic Stability. *Journal of Medicinal Chemistry* **1998**, *41* (27), 5353–5361. <https://doi.org/10.1021/JM970257G>.
- (70) BF, T.; IB, A.; SW, M.; BR, M.; RK, R. Structure-Activity Analysis of Anandamide Analogs: Relationship to a Cannabinoid Pharmacophore. *Journal of medicinal chemistry* **1996**, *39* (2), 471–479. <https://doi.org/10.1021/JM9505167>.
- (71) Richard Silver, by T.; Gordon Douglas, R.; Marcus, A. J.; York, N. COMBINED CLINICAL AND BASIC SCIENCE SEMINAR Selected and Edited Eicosanoids as Bioregulators in Clinical Medicine.
- (72) Willis, A. L. Synthesis of Eicosanoids. **2017**, 1–18. <https://doi.org/10.1201/9780203712634-2>.
- (73) di Marzo, V. New Approaches and Challenges to Targeting the Endocannabinoid System. *Nature Reviews Drug Discovery* **2018** *17*:9 **2018**, *17* (9), 623–639. <https://doi.org/10.1038/nrd.2018.115>.
- (74) Kenakin, T. Functional Selectivity and Biased Receptor Signaling. *Journal of Pharmacology and Experimental Therapeutics* **2011**, *336* (2), 296–302. <https://doi.org/10.1124/JPET.110.173948>.
- (75) T, K.; LJ, M. Seven Transmembrane Receptors as Shapeshifting Proteins:

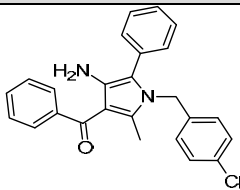
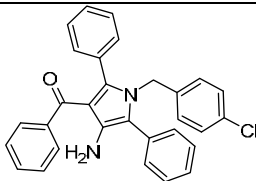
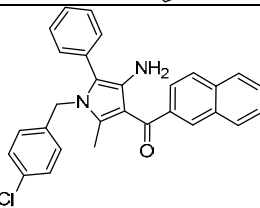
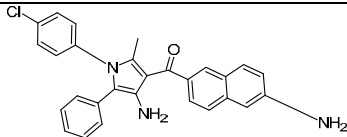
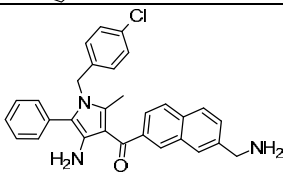
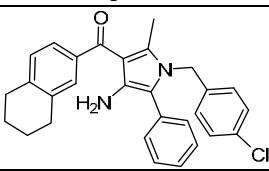
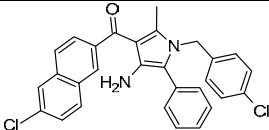
- The Impact of Allosteric Modulation and Functional Selectivity on New Drug Discovery. *Pharmacological Reviews* **2010**, *62* (2), 265–304. <https://doi.org/10.1124/PR.108.000992>.
- (76) Priestley, R.; Glass, M.; Kendall, D. Functional Selectivity at Cannabinoid Receptors. *Advances in Pharmacology* **2017**, *80*, 207–221. <https://doi.org/10.1016/BS.APHA.2017.03.005>.
- (77) Mukhopadhyay, S.; Shim, J. Y.; Assi, A. A.; Norford, D.; Howlett, A. C. CB1 Cannabinoid Receptor-G Protein Association: A Possible Mechanism for Differential Signaling. *Chemistry and Physics of Lipids* **2002**, *121* (1–2), 91–109. [https://doi.org/10.1016/S0009-3084\(02\)00153-6](https://doi.org/10.1016/S0009-3084(02)00153-6).
- (78) Mukhopadhyay, S.; Howlett, A. C. Chemically Distinct Ligands Promote Differential CB1 Cannabinoid Receptor-Gi Protein Interactions. *Molecular Pharmacology* **2005**, *67* (6), 2016–2024. <https://doi.org/10.1124/MOL.104.003558>.
- (79) Prather, P. L.; Martin, N. A.; Breivogel, C. S.; Childers, S. R. Activation of Cannabinoid Receptors in Rat Brain by WIN 55212-2 Produces Coupling to Multiple G Protein α -Subunits with Different Potencies. *Molecular Pharmacology* **2000**, *57* (5), 1000.
- (80) Abrams, D.; Guzman, M. Cannabis in Cancer Care. *Clinical Pharmacology & Therapeutics* **2015**, *97* (6), 575–586. <https://doi.org/10.1002/CPT.108>.
- (81) Moreira, F. A.; Grieb, M.; Lutz, B. Central Side-Effects of Therapies Based on CB 1 Cannabinoid Receptor Agonists and Antagonists: Focus on Anxiety and Depression. *Best Practice & Research Clinical Endocrinology & Metabolism* **23**, 133–144. <https://doi.org/10.1016/j.beem.2008.09.003>.
- (82) Vučković, S.; Srebro, D.; Vujović, K. S.; Vučetić, Č.; Prostran, M. Cannabinoids and Pain: New Insights From Old Molecules. *Frontiers in Pharmacology* **2018**, *0* (NOV), 1259. <https://doi.org/10.3389/FPHAR.2018.01259>.
- (83) Beyer, C. E.; Dwyer, J. M.; Piesla, M. J.; Platt, B. J.; Shen, R.; Rahman, Z.; Chan, K.; Manners, M. T.; Samad, T. A.; Kennedy, J. D.; Bingham, B.; Whiteside, G. T. Depression-like Phenotype Following Chronic CB1 Receptor Antagonism. *Neurobiology of Disease* **2010**, *39* (2), 148–155. <https://doi.org/10.1016/J.NBD.2010.03.020>.
- (84) S, H.; J, T.; DJ, B.; RM, J. Therapeutic Utility of Cannabinoid Receptor Type 2 (CB(2)) Selective Agonists. *Journal of Medicinal Chemistry* **2013**, *56* (21), 8224–8256. <https://doi.org/10.1021/JM4005626>.
- (85) BK, A.; J, W.-M.; C, H.; A, S.; K, M. Functional Selectivity in CB(2) Cannabinoid Receptor Signaling and Regulation: Implications for the Therapeutic Potential of CB(2) Ligands. *Molecular Pharmacology* **2011**, *81* (2), 250–263. <https://doi.org/10.1124/MOL.111.074013>.
- (86) A, D.; K, M. CB2 Cannabinoid Receptors as a Therapeutic Target-What Does the Future Hold? *Molecular Pharmacology* **2014**, *86* (4), 430–437. <https://doi.org/10.1124/MOL.114.094649>.
- (87) Ashton, J. C.; Wright, J. L.; Mcpartland, J. M.; Tyndall, J. D. A. Cannabinoid CB1 and CB2 Receptor Ligand Specificity and the Development of CB2-Selective Agonists. *Current Medicinal Chemistry* **2008**, *15*, 1428–1443.
- (88) Costa, B. On the Pharmacological Properties of Δ 9-Tetrahydrocannabinol

- (THC). *Chemistry & Biodiversity* **2007**, *4* (8), 1664–1677. <https://doi.org/10.1002/CBDV.200790146>.
- (89) Abadji, V.; Lin, S.; Taha, G.; Griffin, G.; Stevenson, L. A.; Pertwee, R. G.; Makriyannis, A. (R)-Methanandamide: A Chiral Novel Anandamide Possessing Higher Potency and Metabolic Stability. *Journal of Medicinal Chemistry* **2002**, *37* (12), 1889–1893. <https://doi.org/10.1021/JM00038A020>.
- (90) McLaughlin, P. J.; Lu, D.; Winston, K. M.; Thakur, G.; Swezey, L. A.; Makriyannis, A.; Salamone, J. D. Behavioral Effects of the Novel Cannabinoid Full Agonist AM 411. *Pharmacology Biochemistry and Behavior* **2005**, *81* (1), 78–88. <https://doi.org/10.1016/J.PBB.2005.02.005>.
- (91) Ogawa, L. M.; Burford, N. T.; Liao, Y.-H.; Scott, C. E.; Hine, A. M.; Dowling, C.; Chin, J.; Power, M.; Hunnicutt, E. J.; Jr.; Emerick, V. L.; Banks, M.; Zhang, L.; Gerritz, S. W.; Alt, A.; Kendall, D. A. Discovery of Selective Cannabinoid CB2 Receptor Agonists by High-Throughput Screening. *SLAS discovery: advancing life sciences R & D* **2018**, *23* (4), 375. <https://doi.org/10.1177/2472555217748403>.
- (92) Adams, J.; Unett, D.; Anthony, T.; Gatlin, J.; Gaidarov, I. APD371: A Potent, Highly Selective, Full Agonist of the Human CB2 Receptor with Sustained Analgesic Effects in Rodents. *The Journal of Pain* **2018**, *19* (3), S2. <https://doi.org/10.1016/J.JPAIN.2017.12.014>.
- (93) Porter, A. C.; Sauer, J.-M.; Knierman, M. D.; Becker, G. W.; Berna, M. J.; Bao, J.; Nomikos, G. G.; Carter, P.; Bymaster, F. P.; Leese, A. B.; Felder, C. C. Characterization of a Novel Endocannabinoid, Virodhamine, with Antagonist Activity at the CB1 Receptor. *Journal of Pharmacology and Experimental Therapeutics* **2002**, *301* (3), 1020–1024. <https://doi.org/10.1124/JPET.301.3.1020>.
- (94) H, S.; L, C.-B.; C, M.; SN, P.; A, K.; ME, A. The Endocannabinoids Anandamide and Virodhamine Modulate the Activity of the Candidate Cannabinoid Receptor GPR55. *Journal of neuroimmune pharmacology: the official journal of the Society on NeuroImmune Pharmacology* **2012**, *7* (4), 856–865. <https://doi.org/10.1007/S11481-012-9351-6>.
- (95) Carnevale, L. N.; Arango, A. S.; Arnold, W. R.; Tajkhorshid, E.; Das, A. Endocannabinoid Virodhamine Is an Endogenous Inhibitor of Human Cardiovascular CYP2J2 Epoxygenase. *Biochemistry* **2018**, *57*(46), 6489–6499. <https://doi.org/10.1021/ACS.BIOCHEM.8B00691>.
- (96) Kozłowska, H.; Baranowska, M.; Schlicker, E.; Kozłowski, M.; Ludański, J.; Malinowska, B. Virodhamine Relaxes the Human Pulmonary Artery through the Endothelial Cannabinoid Receptor and Indirectly through a COX Product. *British Journal of Pharmacology* **2008**, *155* (7), 1034–1042. <https://doi.org/10.1038/BJP.2008.371>.
- (97) J, L.; A, D.; A, T.; G, S.; M, M.; S, R.; N, S.; C, X.; G, T.; D, P. Synthesis and Characterization of a Peripherally Restricted CB1 Cannabinoid Antagonist, URB447, That Reduces Feeding and Body-Weight Gain in Mice. *Bioorganic & medicinal chemistry letters* **2009**, *19* (3), 639–643. <https://doi.org/10.1016/J.BMCL.2008.12.059>.
- (98) Carloni, S.; Crinelli, R.; Palma, L.; Álvarez, F. J.; Piomelli, D.; Duranti, A.;

- Balduini, W.; Alonso-Alconada, D. The Synthetic Cannabinoid URB447 Reduces Brain Injury and the Associated White Matter Demyelination after Hypoxia-Ischemia in Neonatal Rats. *ACS Chemical Neuroscience* **2020**, *11* (9), 1291–1299. <https://doi.org/10.1021/ACSCHEMNEURO.0C00047>.
- (99) Dhopeswarkar, A.; Murataeva, N.; Makriyannis, A.; Straiker, A.; Mackie, K. Two Janus Cannabinoids That Are Both CB2 Agonists and CB1 Antagonists. *Journal of Pharmacology and Experimental Therapeutics* **2017**, *360* (2), 300–311. <https://doi.org/10.1124/JPET.116.236539>.
- (100) Li, A.-L.; Carey, L. M.; Mackie, K.; Hohmann, A. G. Cannabinoid CB2 Agonist GW405833 Suppresses Inflammatory and Neuropathic Pain through a CB1 Mechanism That Is Independent of CB2 Receptors in Mice. *Journal of Pharmacology and Experimental Therapeutics* **2017**, *362* (2), 296–305. <https://doi.org/10.1124/JPET.117.241901>.
- (101) Li, A.-L.; Lin, X.; Dhopeswarkar, A. S.; Thomaz, A. C.; Carey, L. M.; Liu, Y.; Nikas, S. P.; Makriyannis, A.; Mackie, K.; Hohmann, A. G. Cannabinoid CB2 Agonist AM1710 Differentially Suppresses Distinct Pathological Pain States and Attenuates Morphine Tolerance and Withdrawal. *Molecular Pharmacology* **2019**, *95* (2), 155–168. <https://doi.org/10.1124/MOL.118.113233>.
- (102) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *Journal of Cheminformatics* **2012**, *4* (1), 17. <https://doi.org/10.1186/1758-2946-4-17>.
- (103) EF, P.; TD, G.; CC, H.; EC, M.; GS, C.; TI, C.; JH, M.; TE, F. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein science : a publication of the Protein Society* **2021**, *30* (1), 70–82. <https://doi.org/10.1002/PRO.3943>.
- (104) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of computational chemistry* **2009**, *30* (16), 2785. <https://doi.org/10.1002/JCC.21256>.
- (105) Molinspiration Team. Molinspiration Cheminformatics Free Web Services. Slovsky Grob June 2020.
- (106) Skancke, P. N. Divinylketone: Conformational Preferences. Molecular Geometry and Vibrational Force Field in the Ground State. *Journal of Computational Chemistry* **1983**, *4* (2), 142–149. <https://doi.org/10.1002/JCC.540040204>.

Supplementary Materials

Table S1. A concise view of all structures docked in this study, with binding affinity to CB1 and CB2, ratio of CB1 and CB2 affinity, 2D structure, and LogS.

Cpd Name	CB1 (Autodock binding score)	CB2 (Autodock binding score)	CB1/C B2	2D Structure	logS
URB447	313 ± 72 (-10.6)	41 ± 23 (-11.2)	0.94		-8.243
1 (Compound 1)	-12.0	-11.6	1.03		-10.13
2 (Compound 2)	-12.0	-12.7	0.94		-9.648
3	-10.9	-11.7	0.93		-9.517
4	-10.2	-12.7	0.80		-9.363
5	-11.7	-13.0	0.90		-9.639
6	-10.6	-13.6	0.78		-10.14

Cpd Name	CB1 (Autodock binding score)	CB2 (Autodock binding score)	CB1/C B2	2D Structure	logS
7	-9.9	-13.7	0.72		-8.487
8 (SWS002)	-8.8	-12.8	0.68		-9.051
9	-10.3	-11.7	0.88		-9.236
10	-10.6	-11.1	0.95		-10.08
11	-10.8	-10.5	1.03		-11.01
12 (Compound 3)	-12.5	-12.1	1.03		-11.01
13 (SWS001)	-11.8	-11.4	1.04		-8.025



Exploring the Adverse Health Effects Associated with Atrazine Exposure and Whether More Substantial Restrictions Are Required in the United States

Simratpal Deol

Author Background: *Simratpal Deol grew up in the United States and currently attends Ithaca Senior High School in Ithaca, New York in the United States. His Pioneer research concentration was in the field of environmental studies and titled “Environmental Challenges: Environmental Choices.”*

Abstract

Atrazine is the second most applied pesticide in the world, with annual consumption varying from 70,000 to 90,000 tons. It is primarily used on maize, sorghum, and sugarcane, and it is generally effective in combating a range of broadleaf (dicot) and grassy weeds.

Atrazine exposure may be associated with thyroid and ovarian cancers in certified applicators, as well as intrauterine growth retardation, small-for-gestational-age births, premature births, low birth weights, and spontaneous abortions in pregnant women. Women who drink atrazine-contaminated tap water may face increased risk of menstrual irregularity and abnormal hormone levels. Evidence has also tied atrazine exposure, as well as numerous other pesticides, to consistent declines in male sperm quality.

Italy and Germany banned atrazine in 1991 following growing concerns over the contamination of drinking water with pesticides. Finland and Denmark followed later in the 1990s, and the European Union officially prohibited the substance in 2004, citing concerns over its potential role as a carcinogen. Atrazine remains in wide use in the United States, where over 30,000 tons are used yearly.

The atrazine bans in Italy and Germany failed to stifle the nations' maize and sorghum industries, suggesting that the pesticide is not essential in maintaining high agricultural yields. An economic simulation found that corn and sorghum yields in the Midwest would only decrease by 2.4% and 3.7%, respectively, with the implementation of a total atrazine ban in the United States. Prices would increase by merely 1.8% for corn and 2.4% for

sorghum. Given the rising need for integrated weed management strategies in promoting a sustainable future for agriculture, I recommend that the United States implements a total ban on atrazine or, at the very least, more substantial restrictions on the substance.

1. Atrazine: A General Background

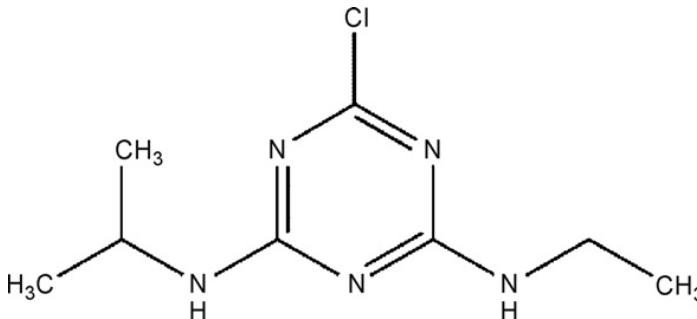


Figure 1. Molecular structure of atrazine

Atrazine (Figure 1), or “6-chloro-n-ethyl-n-(1-methylethyl)-1,3,5-triazine-2,4-diamine,” is a synthetic triazine herbicide (Singh et al., 2018) that was first registered in the United States by the manufacturer CIBA-GEIGY in 1958 (Ribaud and Bouzahr, 1994). Atrazine is popular due to its comparatively low application cost and season-long effects (Swanton et al., 2007). The annual global consumption of atrazine is estimated to be between 70,000 and 90,000 tons, making it the second most applied pesticide in the world (Singh et al., 2018). In the United States, where between 30,000 and 34,000 tons of the substance is applied every year (Cole, 2012), atrazine only trails glyphosate (Roundup) in terms of usage (Hites and Raff, 2013).

Atrazine has historically been effective in suppressing broadleaf (dicot) weeds, including redroot and green pigweed, cocklebur, wild mustard, lambsquarters, wild buckwheat, common ragweed, and velvet leaf, as well as some grassy weeds (Ribaud and Bouzahr, 1994; Swanton et al., 2007). The vast majority (98%) of atrazine applied in the United States is used on corn, sorghum, and sugarcane fields, though it is also dispensed on wheat, conifer, nut, banana, pineapple, asparagus, guava, and coffee crops, among many others (Little, 1992; Singh et al., 2018). The herbicide is also applied in residential areas in small quantities (Little, 1992). Only certified applicators can administer atrazine in the United States (Freeman et al., 2011).

In agricultural settings, atrazine is typically applied either before any crops are planted or before they emerge. However, atrazine’s selectivity in targeting plants allows farmers to use it after the emergence of crops as well (Ribaud and Bouzahr, 1994; Swanton et al., 2007). Atrazine can be used on its own or in conjunction with other herbicides, such as glufosinate, metolachlor, S-metolachlor, mesotrione, and isoxaflutole (Swanton et al., 2007; Krämer et al., 2012). The increasing frequency with which farmers pair atrazine with other treatments may partially explain why the per-acre usage of atrazine on

cornfields has dropped in both the United States, where application rates declined from 1.50 lbs/acre between 1966 and 1982 to 1.12 in 1992 (Ribaud and Bouzahr, 1994), and Canada, which has seen a decrease from 4.02 lbs/acre in the early 1970s to between 0.45 to 1.47 in 2004 (Swanton et al., 2007).

Like other triazine herbicides, atrazine is classified as a selective photosystem II inhibitor (Swanton et al., 2007) due to its ability to block electron transport in the chloroplasts of target plants (Krämer et al., 2012). Actively developing crops such as corn and sorghum can metabolize and deactivate atrazine once it is absorbed (Ribaud and Bouzahr, 1994), preventing the herbicide from seriously damaging agricultural yields. Atrazine exhibits a half-life of between 60 and 150 days in loamy soils and about 660 days in anaerobic sandy clay soils (Ribaud and Bouzahr, 1994).

With a soil absorption coefficient (KOC) ranging from 39 to 13,600 (Little, 1992), atrazine is considerably mobile in soils and is not absorbed in great quantities (Singh et al., 2018), allowing it to run off into the surrounding environment. About 1% of applied atrazine is generally washed into nearby surface waters (Freeman et al., 2011), though some estimates place the total runoff percentage as high as 18% (Little, 1992). Atmospheric transport and precipitation can deposit atrazine over 1,000 kilometers away from where it was initially applied (Freeman et al., 2011).

A growing subset of scientists, national governments, and international agencies have expressed concerns over atrazine's long-term consequences on both humans and the natural world. In 2002, a study on wild leopard frogs (*Rana pipiens*) throughout the United States found that populations exposed to minimally detectable doses of atrazine contained males with testicular oocytes and other evidence of hermaphroditism. At one site in Wyoming's North Platte River, which is fed by numerous streams from atrazine-intensive areas, 92% of male wild leopard frogs showed evidence of sex reversal. Further laboratory experiments found that testicular oocytes were found in 29% of frogs exposed to 0.1 atrazine parts per billion (p.p.b.), compared to 0% in the unexposed control (Hayes et al., 2003). Tyrone B. Hayes, the lead author of the study, has since become involved in a lengthy legal battle with Syngenta, the primary manufacturer of atrazine, over the safety of the herbicide (Aviv, 2014).

Italy, Germany, Denmark, and Finland banned atrazine in the 1990s (Singh et al., 2018), and the remainder of the European Union barred its application in 2004, citing the possibility that the herbicide may be linked to cancer in humans (Krämer et al., 2012). Though atrazine use is permitted in the United States, the nation's Environmental Protection Agency (EPA) places bodies of water with over 3 µg atrazine/L in its Atrazine Monitoring Program (Almberg et al., 2018). The organization classifies atrazine as an "endocrine disruptor," and the International Agency for Research on Cancer (IARC) lists it as a "carcinogenic pesticide" (Singh et al., 2018).

2. Potential Risks to Humans

Given the biological differences between wild animal species and humans, findings on the effects of atrazine on wild leopard frogs and other amphibians cannot be reliably extrapolated to humans (Freeman et al., 2011; Cragin et al., 2011). Nonetheless, studies conducted on human populations have found

associations between atrazine exposure and ovarian cancer (Little, 1992), thyroid cancer (Freeman et al., 2011), low birth weight in newborns (Almberg et al., 2018), preterm births, births that are small-for-gestational-age (Villanueva et al., 2005), intrauterine growth retardation (Munger et al., 1997), menstrual cycle length irregularity, longer follicular phases, reduced levels of hormones involved in the menstrual cycle (Cragin et al., 2011), and reductions in sperm quality and production (Sengupta and Banerjee, 2013).

In further exploring the effects of atrazine on humans, I will primarily focus on research involving (1) certified atrazine applicators and their immediate families and (2) representative samples of humans across all occupations living in certain loci. Because applicators and their immediate families are exposed to higher doses of atrazine than the average human, they may be predisposed to harsher illnesses, should they exist. Those working with atrazine in occupational settings are primarily exposed to the herbicide via dermal absorption and/or inhalation, while the general population is primarily exposed to the herbicide via contaminated drinking water (Villanueva et al., 2005).

2.1 Cancer Incidence in Certified Atrazine Applicators in the Agricultural Health Study

The Agricultural Health Study (AHS) is a prospective cohort study jointly managed by the National Institute of Environmental Health Sciences (NIEHS), EPA, National Institute for Occupational Safety and Health (NIOSH), and National Cancer Institute (NCI). The AHS includes 52,395 private pesticide applicators, primarily farmers, from North Carolina and Iowa, 4,916 commercial applicators from Iowa, and 32,347 spouses who are married to applicators in the study (NIEHS, 2019). Applicators were enrolled in the AHS between 1993 and 1997 when receiving or renewing licenses for restricted use pesticides (Freeman et al., 2011). Upon enrollment, applicators completed a questionnaire in which they described the duration (number of exposed years) and frequency (average number of days used per year) of their past exposure to 22 pesticides (including atrazine), nine insecticides, two fungicides, and one fumigant (Dosemeci et al., 2002).

Those who completed the initial set of questions received additional take-home questionnaires in which they described their use of 50 agricultural compounds, methods for pesticide application and mixing, repair strategies for application equipment, and use of protective equipment when working with pesticides (Freeman et al., 2011; National Institutes of Health (NIH), 2019). Participants also disclosed personal information, including their age, sex, level of education, past medical history, familial cancer history, diet, alcohol use, and smoking habits (Freeman et al., 2011). The featured compounds were chosen because of their importance to agriculture in Iowa and North Carolina and/or evidence that they may be associated with adverse health effects in humans or animals (Dosemeci et al., 2002). Data from the questionnaires were used to calculate (1) each applicator's lifetime days of atrazine use and (2) each applicator's intensity-weighted lifetime days of atrazine use (Dosemeci et al., 2002; Rusiecki et al., 2004; Freeman et al., 2011). The formulas for these measurements are shown as follows:

Table 1. *Formulas for calculating lifetime days of atrazine use, intensity-weighted lifetime days of atrazine use, and intensity level, which is a part of the formula for intensity-weighted lifetime days of use (Dosemeci et al., 2002; Rusiecki et al., 2004).*

<p>Lifetime Days of Atrazine Use = (Reported Years of Use) x (Average Days of Use Per Year)</p>
<p>Intensity-Weighted Lifetime Days of Use = (Lifetime Days of Atrazine Use) x (Intensity Level*)</p>
<p>*Intensity Level = ((Mix x Enclosed) + (Appl x Cab) + Repair + Wash) x PPE x Repl x Hyg x Spill</p> <p><u>VARIABLE DEFINITIONS:</u> Mix = Status of Pesticide Mixing Enclosed = Whether Enclosed Mixing System Was Used Appl = Application Methods Cab = Tractor with Enclosed Cab and/or Charcoal Filter Repair = Status of Repairing Equipment Wash = Status of Washing Equipment After Application PPE = Use of Personal Protective Equipment During Application Repl = Replacement of Gloves Hyg = Personal Hygiene (e.g., washing hands, taking showers after applying pesticides) Spill = Changing Clothes After Atrazine Spill</p>

Freeman et al. (2011) compared the rate of cancer development in atrazine applicators enrolled in the AHS to that of AHS applicators who did not report using atrazine. Of the 57,311 applicators initially included in the AHS, the authors excluded over 3,500 because of previous cancer diagnoses and/or missing data. 36,357 of the remaining 53,662 applicators reported using atrazine in the provided questionnaires, of whom 3,146 (8.65%) developed cancer between enrollment and December 31, 2007. That figure alone was, from a statistical standpoint, not significantly different from the 1,591 (9.19%) non-atrazine applicators who were diagnosed with cancer over a similar time period.

Freeman et al. (2011) further analyzed the specific cancers reported in over 20 AHS atrazine applicators. The authors also divided the sample of atrazine applicators into quartiles based on their use of the herbicide, as dictated by lifetime days or intensity-weighted lifetime days. Via Poisson regression models, they calculated the relative risks (RRs) carried within each group and whether certain cancers were significantly more prevalent in atrazine applicators. Despite their infrequent occurrences amongst included applicators, ovarian and breast cancer were also included in the analysis due to existing concerns about atrazine's hormonal tendencies.

Based on 29 cases among the atrazine applicators, a statistically significant association between thyroid cancer and intensity-weighted lifetime days of atrazine use was found in the highest quartile of users. Even when adjusting the statistical model to account for body mass index (past studies have linked obesity to increased risk for thyroid cancer), the thyroid cancer risk for applicators in the second and fourth quartiles of intensity-weighted use was

significantly higher than those in the lowest quartile. Though a statistically significant association was found between thyroid cancer and top-quartile lifetime days of atrazine use, accounting for obesity eliminated this pattern (Freeman et al., 2011).

Freeman et al. (2011) also observed a non-statistically significant increase in the risk of ovarian cancer in women who directly applied atrazine, though the low number of female atrazine users (312) in the AHS prevented the authors from evaluating a potential exposure-response relationship. Past studies using data from the AHS have suggested that pesticide use is associated with the development of ovarian cancer; through 2002, the 1,563 female applicators in the AHS showed increased rates of ovarian cancer development in comparison to both the general population and the 32,127 female spouses of male AHS applicators (Waggoner et al., 2010). Donna et al. (1989) found that women who were previously exposed to triazines—the class of pesticides that contains atrazine—were 2-3 times more likely to develop epithelial ovarian cancer compared to women never exposed.

Freeman et al. (2011) did not report any noteworthy links between atrazine exposure and other types of cancer. An analysis of AHS data through 2001 had previously found non-significant but suggestive associations between atrazine use and non-Hodgkin's lymphoma, multiple myeloma, bladder cancer, and lung cancer (Rusiecki, 2004). Case-control studies outside the AHS have also shown possible links with non-Hodgkin's lymphoma (De Roos et al., 2003), as well as prostate cancer (MacLennan et al. 2002). Additional data are needed before clearer relationships can be established.

2.2 Adverse Effects on the Female Reproductive System

Atrazine exposure has been linked to numerous adverse effects on the female human reproductive system, such as intrauterine growth retardation, small-for-gestational-age births, premature births, low weight at birth, and spontaneous abortion (Munger et al., 1997; Villanueva et al., 2005; Cragin et al., 2011; Freeman et al., 2011). Female pesticide applicators, including those who applied atrazine, in the AHS showed an increased risk of intermenstrual bleeding and missed periods (Farr et al., 2004). Irregularities in the menstrual cycle are, in turn, associated with various cancers, osteoporosis, cardiovascular disease, diabetes, and other hormonally influenced conditions (Cragin et al., 2011), making atrazine's effects on the female reproductive system important to dissect.

One study examined the relationship between atrazine exposure from drinking water and patterns in the menstrual cycle. The 102 participants were premenopausal Illinois or Vermont women between 18 and 40 years of age who were not on hormone-influencing medications and/or birth control, had not reported a pregnancy in the past six months, had not documented breastfeeding in the past three months, and had no known history of endocrine and/or reproductive disorders. The Illinois women were chosen from Mount Olive and Gillespie, two communities with some of the highest concentrations of atrazine in municipal drinking water in the entire nation. The Vermont women were chosen from communities where atrazine is used sparingly (Cragin et al., 2011).

All 102 women completed a questionnaire on their reproductive history,

their menstrual cycle characteristics, possible confounding variables (e.g., age, race, education, caffeine intake, etc.), and exposure indices. 67 women provided a daily diary of vaginal bleeding for a complete menstrual cycle, as well as paired urine and tap water samples from a single day. 35 tap water samples, 15 from Illinois and 20 from Vermont, contained detectable levels of atrazine and general chlorotriazines. The Illinois samples contained an average atrazine concentration of 0.7 µg/L, nearly double what was found in the 20 samples (0.3 µg/L) from Vermont. None of the Vermont samples exceeded atrazine concentrations of 0.5 µg/L, while all 15 of the Illinois samples passed this threshold (Cragin et al., 2011).

Illinois women were significantly more likely to go more than six weeks without menstruation compared to their Vermont counterparts. Cycle length irregularity was, to a statistically significant extent, also more frequently reported in Illinois. Among Illinois women only, abnormally long cycles were more prevalent among those consuming unfiltered tap water. Atrazine exposure was significantly associated with reduced mid-luteal phase levels of E13G, a metabolite responsible for circulating the primary estrogen in humans, estradiol. Reductions in estrogen circulation have been linked to compromised implantation, reduced overall fertility, and increased risk of osteoporosis, cardiovascular disease, and disorders in the central nervous system. Atrazine use was also connected to reduced mid-luteal phase Pd3G levels, a deficiency that decreases both progesterone secretion and luteal function. These reductions, in turn, may lead to repeated spontaneous abortions, infertility, and compromised implantation (Cragin et al., 2011).

Because of an ongoing drought, the atrazine concentrations in the Mount Olive and Gillespie water samples were atypically low in the year of the study. Only 43% of the collected tap water samples contained atrazine levels over the established detection limit of 0.5 µg/L, and only 5% of participants had measurable atrazine levels in their urine. The potentially serious adverse effects that atrazine has at such low quantities may suggest that the EPA's current threshold of 3 µg/L is too lenient (Cragin et al., 2011).

A separate study conducted in Finistère, France, sought to evaluate atrazine's potential effects on preterm delivery, small-for-gestational-age birth, and low birth weight. Finistère is a predominantly agricultural administrative district—280 tons of pesticides, including 71 tons of atrazine, were used on maize in the region in 1997. The authors used birth records from local government offices to determine the frequency of adverse birth effects in Finistère between October 1, 1997, and September 30, 1998. In addition to this, the researchers obtained data on atrazine levels in municipal water sources between 1990 and 1998 from the Direction Départementale des Affaires Sanitaires et Sociales (which translates to the "District Health and Social Affairs Bureau") and divided the included communities in tertiles based on atrazine concentration. The atrazine concentrations present at each location were compared to the number of adverse birth occurrences in those areas using odds ratios (ORs) and 95% confidence intervals (CIs) (Villanueva et al., 2005).

Infants featured in the Finistère study were deemed "low weight" if they weighed less than 2,500 grams at birth, "preterm" if delivered at less than 37 weeks, and "small-for-gestational-age" if below the tenth percentile for birth weight at their gestational age. Of the 3,510 newborns included, 163 were low-

weight births, 137 were premature, and 241 were small-for-gestational-age. The researchers found that the risk of small-for-gestational-age births was significantly elevated if the third trimester occurred partially or completely between May and September, when atrazine concentrations in municipal water sources were consistently highest due to heightened agricultural activity. The risk of prematurity was increased at a borderline significant level when the first trimester occurred between May and September, though no dose-response trend was observed between atrazine exposure and preterm births (Villanueva et al., 2005).

Although the possible presence of confounding pollutants (including lead, polychlorinated biphenyls, and alternative pesticides) may have affected the results of the Finistère study, other analyses of atrazine corroborate its findings. An Iowa survey of 856 municipal drinking water sources between 1986 and 1987 found that the Rathbun water system, which is centered around Rathbun Lake in the southern part of the state, contained elevated levels of triazine herbicides. A subsequent study of local singleton births from 1984 to 1990 found that communities near Rathbun Lake faced a greater risk of intrauterine growth retardation. Though the study could not verify a causal relationship involving a specific triazine herbicide, the association was strongest with atrazine (Munger et al., 1997).

An additional study of 14,445 singleton births in Ohio communities enrolled in the EPA's Atrazine Monitoring Program between 2006 and 2008 found an association between atrazine exposure during the first and second trimesters of pregnancy—as well as the overall gestational period—and low birth weight. Those born at an abnormally low weight exhibit an increased risk of infant mortality, lifetime cardiovascular disease, and all-cause mortality (Wilcox et al., 2010), so such findings may be cause for concern. Since only 4% of the water samples taken during the study exceeded the Atrazine Monitoring Program's threshold of 3 µg/L, current restrictions on atrazine may once again be too lenient to prevent adverse birth effects (Almberg et al., 2018).

2.3 Atrazine's Link to Declining Male Fertility

A 1992 analysis of male reproductive data from 20 countries spanning over five decades found that the mean concentration of sperm in male ejaculate had, by a significant margin, decreased from 113 million sperm cells/mL of semen in 1940 to 66 million sperm cells/mL in 1990. A substantial decrease in mean seminal volume—from 3.40 mL to 2.75 mL—was also observed over the same time span. The study included 61 papers involving 14,947 men between the ages of 17 and 64, with the mean age of 30.8 remaining constant throughout the 5-decade period. Similar decreases in sperm density and seminal volume were observed when the authors only included the 28 reports from the United States (Carlsen et al., 1992).

The overall trend of declining semen quality has coincided with a twofold-to-fourfold increase in testicular cancer rates, as well as cryptorchidism—the failure of the testes to fully descend from the abdomen—and hypospadias—a condition where the urethral opening is on the underside of the penis (Carlsen et al., 1992). There may be an association between testicular cancer and lower sperm counts: a separate study found that Danish men, a

demographic diagnosed with testicular cancer at a rate five times higher than Finnish men, have close to half the sperm cell density (70 million sperm cells/mL) of Finnish men (131 million sperm cells/mL) (Nikkanen et al., 1979). Sperm cell density has also been correlated with male fertility, and analyses have suggested that the proportion of infertile men has increased (Bostofte et al., 1982).

When Carlsen et al. (1992) presented their findings to the World Health Organization (WHO), the authors noted that “such remarkable changes in semen quality and the occurrence of genitourinary abnormalities over a relatively short period is more probably due to environmental rather than genetic factors” (p.612). Numerous works have since connected general pesticide use to problems within the male reproductive system. Abell et al. (2000), for example, found that greenhouse workers with more than 10 years of experience had lower sperm concentrations than those with less than five years of experience, perhaps, in part, due to longer periods of exposure to pesticides. In addition, males previously exposed to dichlorodiphenyltrichloroethane (DDT), a powerful insecticide banned by the United States in 1972, may have decreased semen volumes following ejaculation, lower overall sperm counts (Ayotte et al., 2001), and reduced serum testosterone levels (Martin et al., 2002).

The application of atrazine could potentially be related to the observed decline in male fertility. Previous reports have found that atrazine facilitates the production of aromatase, an enzyme that effectively converts androgen, a male sex hormone, to estrogen, specifically estradiol (Sengupta and Banerjee, 2013). Though the exact functioning of this mechanism must be further studied in humans, the replacement of androgens with estradiol is believed to reduce both sperm concentration and overall sperm count in rats (Hayes et al., 2003; Freeman et al., 2011).

A separate experiment found that exposing incubated samples of human semen to increasing concentrations of atrazine decreased the overall motility of sperm cells, a trend perhaps explained by atrazine-induced mitochondrial dysfunction. In their analysis, Hase et al. (2008) proposed that atrazine targets F1FO-ATP synthase, a protein that, if disrupted, interferes with sperm cells' ability to produce ATP for proper movement and function. The same researchers found that semen samples exposed to atrazine possessed significantly less ATP than others; the reduction in ATP synthesis matches what was seen with the oligomycin, a compound known to bind to F1FO-ATP synthase, control, further suggesting that atrazine may directly disrupt cellular processes within sperm cells.

3. Importance of Atrazine to Farmers and the Agricultural Industry

When considering implementing stronger restrictions on certain pesticides, governments must also discuss the economic importance of those pesticides to farmers and the broader agricultural industry. Directly comparing the financial benefits of pesticides to the potential adverse health effects that they may have on wild animals and humans can devolve into abstract and subjective

speculation. To best preserve objectivity, this paper reviews several analyses of agricultural data from the United States that (1) determine the true impact that atrazine has on crop yields, prices, and overall profit in comparison to safer alternatives and/or (2) explore the economic effects that further restrictions on atrazine, including a total atrazine ban, would have on a national scale. The short and long-term implications of past atrazine bans in European nations have also been detailed, as well as the growth of atrazine-resistant weeds.

3.1 Findings from the Comprehensive Environmental Economic Policy Evaluation System

Based in Iowa State University, the Comprehensive Environmental Economic Policy Evaluation System (CEEPES) includes a set of models that simulate the effects of domestic environmental policies, including restrictions on pesticide use, in the United States. The CEEPES algorithm possesses four primary components: policy, which describes proposed legislation and alternative management strategies; agricultural production decision, which includes the specific methods used by individual farmers after legislative changes have been made; fate and transport, which accounts for soil runoff and other damage-relevant concentration measures; and environmental exposure, based on an assigned "exposure value" that is calculated by dividing the predicted concentration of a given pesticide with its established environmental benchmark (Ribaudo and Hurley, 1997).

The CEEPES system tracks a multi-state portion of the Midwest that includes more than 80% of the United States' corn acreage, making it an effective tool to study hypothetical policies on atrazine application. Using CEEPES projections, Ribaudo and Hurley (1997) evaluated the economic implications of (1) an American ban on atrazine use, (2) restrictions that would allow farmers to only use the herbicide after crops emerge, and (3) a separate set of legislative changes in atrazine management strategy wherever the compound is found to contaminate surface water.

Under the simulated ban on atrazine use, corn grain and sorghum yields decreased by 1.2% and 3.4% in the study region, respectively. Compared to the EPA's estimation that an atrazine ban would decrease domestic corn yields by 6%, the CEEPES model is a bit more forgiving. The acreage of corn declined by 2.4% in the studied region, while the acreage of sorghum decreased 3.7%. The overall prices of corn and sorghum increased by 1.8% and 2.4%, respectively, as a result of the decreased crop turnouts, and producer income was reduced by \$268 million (Ribaudo and Hurley, 1997).

The per-acre weed treatment cost for corn increased by 10%, in part because of the low costs traditionally associated with atrazine application. The simulation predicted that alternative herbicides would become more widely used in place of atrazine—a 25.6% increase on the use of non-triazine herbicides on corn was observed, while only a 20.6% decrease in triazine herbicide application was seen (note that atrazine is responsible for 54% of all triazine herbicides used on corn). A 36% increase in per-acre weed treatment cost was observed in sorghum fields, as was an 85.6% rise in the use of non-triazine herbicides. Several of the alternative herbicides used in place of atrazine, such as cyanazine, which is considered a more potent carcinogen than atrazine, and

simazine, had few imposed drinking water standards at the time of the study, and comparatively little research had been done on their potential effects on the environment.

Ribaudo and Hurley (1997) also explored the effects of restricting atrazine use to post-emergent applications. Under these conditions, the researchers found that corn yields dropped by 1.2%, while sorghum yields increased by 0.4%. Furthermore, corn and sorghum acreage decreased by 2.6% and 0.5%, respectively. The costs for weed treatment increased by 5.6% for corn and 17.4% for sorghum, both lower than what was found with a total atrazine ban. The use of non-triazine herbicides increased by a lesser margin: a 21.2% increase was observed in corn, while a 42.7% increase was found in sorghum. The CEEPES model projected that the overall producer income would decline by \$204 million.

Ribaudo and Hurley (1997) claimed that a third environmental strategy (controlling atrazine use in areas where the compound can threaten surface water) would be the most financially efficient. The CEEPES model was implemented to (1) assess the economic implications of restricting the concentration of atrazine in surface water to the 10-day health advisory level (HAL) of 100 p.p.b. and (2) give farmers ample leeway in adjusting their farming methods to fulfill water quality standards. Corn grain yields decreased by 1.2% under these circumstances, while corn silage yields rose by 0.1%. Though sorghum yields decreased 2.8%, the per-acre treatment cost for corn and sorghum increased by comparatively smaller margins—4.6% and 27%, respectively. The researchers also noted that undertaking this strategy would result in the lowest pesticide exposure values out of the methods analyzed, in part because farmers would have to operate under a more controlled set of weed management practices. The price of corn increased by a mere 1.4%, while sorghum prices rose by 1.7%.

The CEEPES' findings on the total atrazine ban were included in a 1994 USDA report about the herbicide and subsequently interpreted by Ackerman (2007). Though the strategy was not considered the most cost-effective approach to regulating atrazine, Ackerman (2007) argued that the potential hazards associated with atrazine use outweighed any minimal financial benefits that an alternative management strategy may provide:

“If, on the other hand, the yield impact is on the order of 1%, as USDA estimated, or close to zero, as suggested by the newer evidence discussed here, then the economic consequences become minimal. The USDA study, with a 1% yield loss, found a slight net economic benefit to producers; the entire economic loss in that study came from the impact on consumers, due to the increase in corn prices of almost 2%. It is hard to believe that [the booming corn] industry could not withstand the—remarkably small—economic impacts of banning atrazine.”
(*Ackerman, 2007, p. 448*)

3.2 Implications of Atrazine Bans in Italy and Germany

Both Italy and Germany banned atrazine in 1991, over a decade before the European Union formally banned the substance (Ackerman, 2007). The Italian government had been facing growing public scrutiny over the safety of its drinking water in the 1980s. Corn, sorghum, and soy growers in Italy used as much as 4 kg active atrazine ingredient/ha of farmland, well over the 2.5 lbs/acre limit currently imposed in the United States. A water sample taken from an aqueduct in the province of Ferrara in 1986 contained 2.09 μg atrazine/L, twentyfold the 0.1 μg /L limit adopted by Italy the year prior. Subsequent ordinances banned residents of the neighboring Po Valley from using tap water for drinking and preparing food. Germany's regulation came five years after a chemical company, now owned by Syngenta, dumped 400 liters of atrazine into the Rhine River, impeding efforts to clean the water (Ackerman, 2007).

Ackerman (2007) accessed data from the UN Food and Agriculture Organization's (FAO's) Corporate Statistical Database (FAOSTAT) to analyze the effects of these bans on nationwide crop growth. To build on his findings, which solely included information on maize through 2001, I used FAOSTAT to compile yearly data on maize yield (see figure 2) and production (see figure 3) in Italy and Germany from 1980 to 2019. As noted by Ackerman (2007), no immediate decline in maize yield and production was observed; in fact, maize yields increased by 30.7% and 37.0% in Italy and Germany, respectively, between 1991 and 2002. Maize production soared by 69.2% in Italy and 93.0% in Germany during the same time period. The steep declines seen in 2003 reflect a heat wave and drought that ravaged crops throughout the European continent (Friend et al., 2005). Researchers speculate that the stagnation seen in maize production in more recent years is at least partially caused by changes in climate, as opposed to further restrictions on pesticides (European Union, 2019).

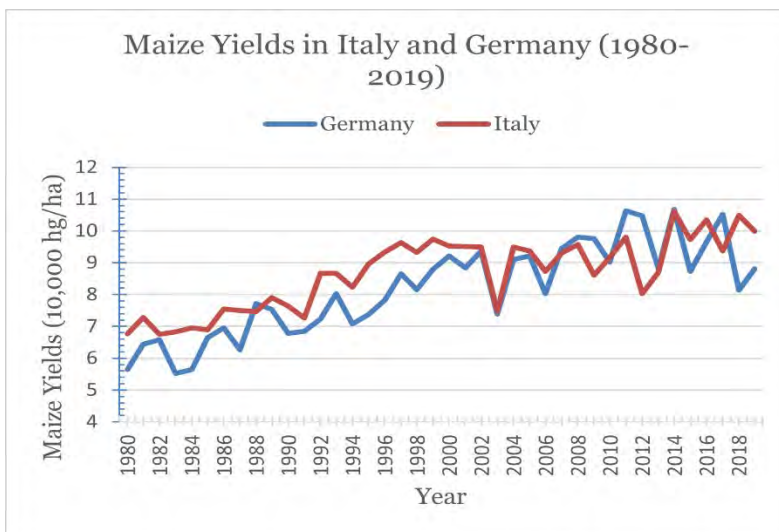


Figure 2. Maize yields in Italy and Germany from 1980 to 2019 (FAO, 2021)

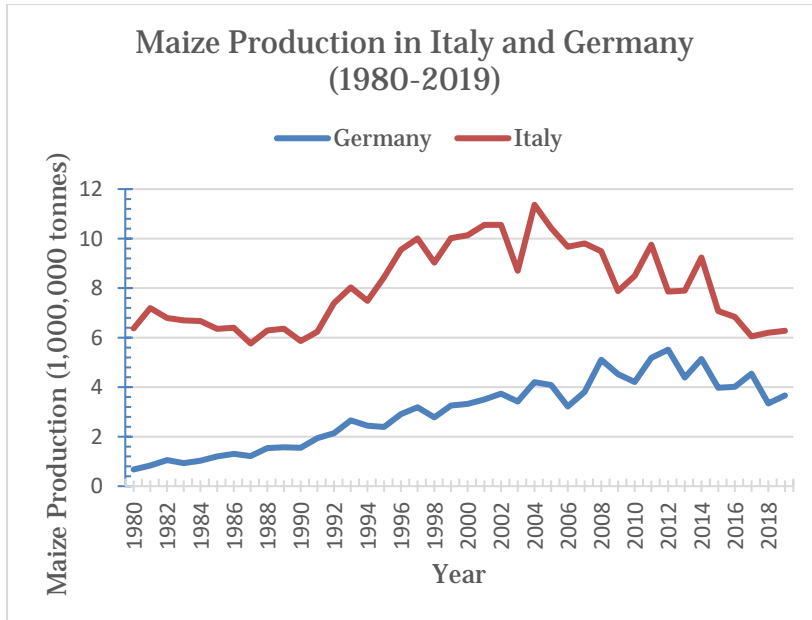


Figure 3. *Maize production in Italy and Germany from 1980 to 2019 (FAO, 2021)*

I also graphed developments in sorghum yield (see figure 4) and production (see figure 5) in Italy—FAOSTAT lacked adequate data on sorghum growth in Germany—between 1980 and 2019. Sorghum yields increased by 14.6% between 1991 and 2002 and have further climbed in the years since, while nationwide sorghum production rose by 43.9% in that same 11-year span and have doubled since 1991. The growth in maize yields in the years following Italy and Germany’s atrazine ban was similar to that of the United States, where atrazine use continued. Moreover, the total area of harvested maize in Italy and Germany increased faster in the 1990s than in the United States, where the metric remained somewhat constant. Given the varying soil and climate conditions in Europe and the United States, the presence of such trends do not necessarily mean that the United States would experience the same growth following a future atrazine ban. However, historical data from Italy and Germany suggest that the prohibition of atrazine failed to significantly impact agricultural development (Ackerman, 2007).

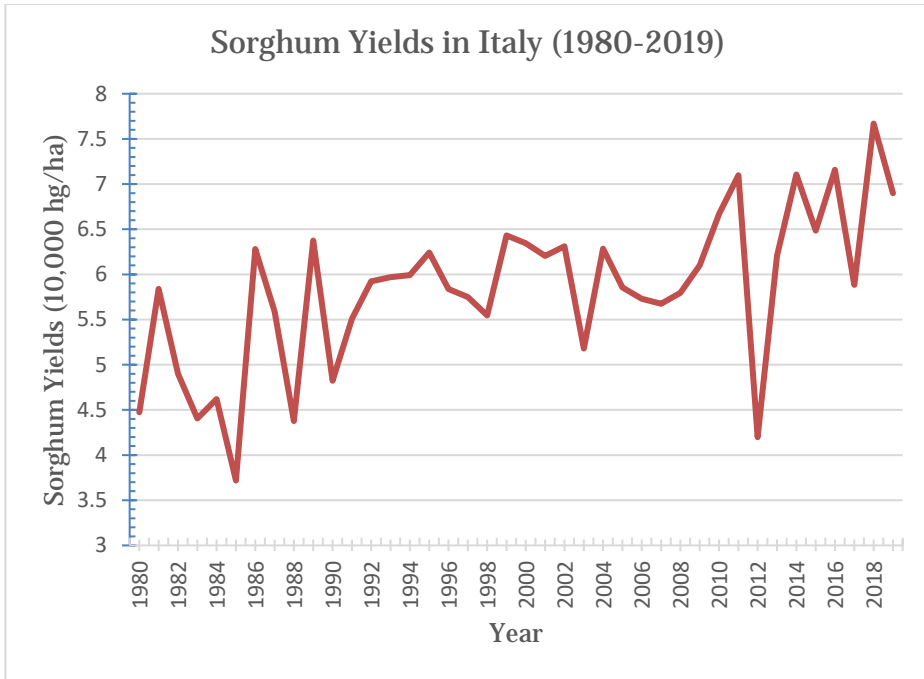


Figure 4. Sorghum yields in Italy from 1980 to 2019 (FAO, 2021)

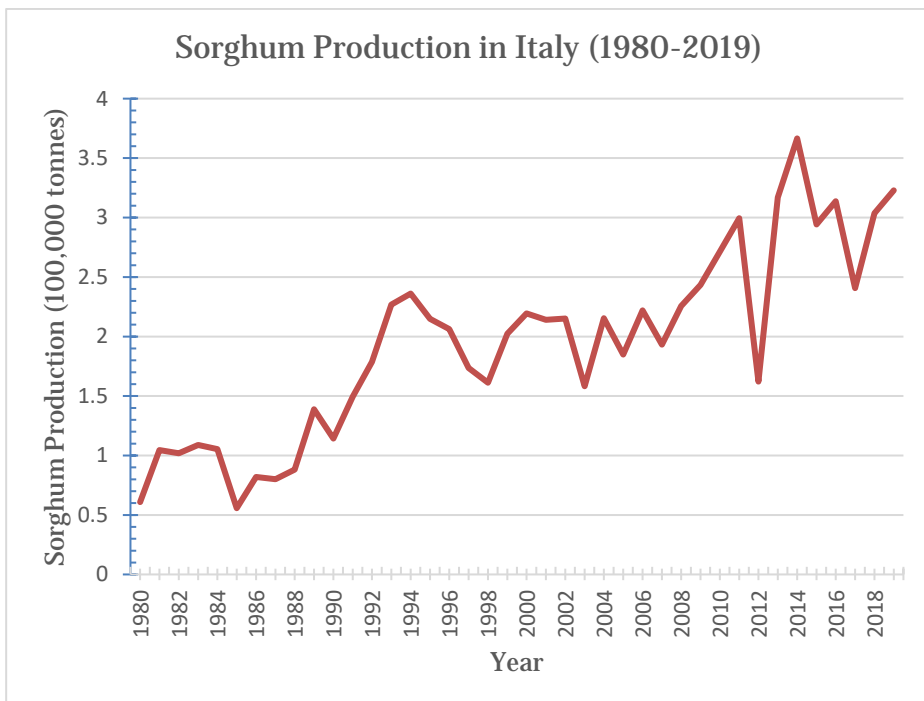


Figure 5. Sorghum production in Italy from 1980 to 2019 (FAO, 2021)

3.3 Resistant Weeds and Alternative Strategies

Though there appears to be widespread agreement within the scientific community that further restrictions on atrazine can be implemented without significant economic downturn, Syngenta, the primary manufacturer of atrazine, has continued to deny claims that the substance warrants additional legislation. In 2011, Syngenta's Atrazine Benefits Team (ABT) released a set of five papers that advocated for future use of the pesticide (Ackerman et al., 2014). One ABT-backed report argued that atrazine is a necessity in the United States because many weeds have developed substantial resistance to numerous other pesticides, including glyphosate (Owen, 2011).

Ackerman et al. (2014) dismissed this claim as superficial, since Owen (2011) made little mention of the possibility of weeds growing resistant to atrazine. In fact, as with many widespread herbicides, weed resistance to atrazine has noticeably expanded over the decades. One analysis found that 10 weeds—foxtails, pigweeds, common waterhemp, common lambsquarters, velvetleaf, other ragweeds not subsequently mentioned, giant ragweed, Palmer pigweed (also known as "Palmer amaranth"), cocklebur, and morning glories—are responsible for 70% of the corn yield loss that can be attributed to weed pressure. Of those 10 weeds, six—foxtails, pigweeds, common waterhemp, common lambsquarters, velvetleaf, and Palmer amaranth—have shown resistance to atrazine within the United States, a trait that was essentially unknown just five decades ago (Ackerman et al., 2014). A separate experiment authored by Michael Owen, the ABT contributor who wrote about atrazine's supposed usefulness in combatting weed resistance, analyzed populations of common waterhemp in Iowa and found that 57% of individuals had developed a resistance to atrazine (Owen, 2013).

Alternative herbicides on the market may provide ample protection to crops in place of atrazine (see tables 2 and 3). When mesotrione, another herbicide produced by Syngenta, was registered in the United States in 2001, an official EPA report predicted that "(mesotrione) will replace atrazine and [isoxaflutole] herbicides" (Ackerman, 2007). A 1999-2001 study funded by Syngenta evaluated the effects of 10 different combinations of herbicides, all of which included mesotrione, atrazine, and/or acetochlor, and found that the most successful treatment that contained mesotrione (but not atrazine) outperformed the most successful treatment that included atrazine by a 2% margin. More recently, Iowa State University's *2013 Herbicide Guide for Iowa Corn and Soybean Production* reported that mesotrione was, with respect to atrazine, comparably successful in combatting the weeds most threatening to corn yields (see tables 2 and 3) (Owen, 2013). Ackerman et al. (2014) described mesotrione as "having modes of action that remain effective against weeds with resistance to atrazine, glyphosate, and ALS inhibitors" (p.63). Saflufenacil and Equip (a combination of foramsulfuron and iodosulfuron) may also serve as viable alternatives in pre-emergent and post-emergent applications, respectively.

Despite atrazine's low cost, replacing it with alternative treatments may be profitable for farmers. A 2008 analysis of weed control profitability in Missouri used WeedSOFT, an economic model that measures the net returns of pest management practices, to compare the financial practicality of hundreds of

herbicide combinations in combatting the state's 10 most prominent weeds (Prato and Woo, 2008). The researchers found that two-pass treatments, in which a pre-emergent and post-emergent herbicide are paired, were consistently the most profitable. Atrazine was present in only 20% of the 70 two-pass treatments deemed profitable by Prato and Woo (2008), and the pre-emergent treatments containing atrazine were markedly less profitable than the two-pass and post-emergent treatments included in the analysis.

Table 2. Effectiveness ratings of pre-emergent herbicides on a series of corn-threatening weeds (1=poor, 2=fair, 3=good, 4=excellent) (Owen, 2013; Ackerman et al., 2014).

	Atrazine	Saflufenacil	Mesotrione	Flumesulam + clopyralid
Foxtail	2	1	1	1
Pigweeds and Waterhemp	4	3.5	3.5	3.5
Lambsquarters	4	3.5	4	3
Velvetleaf	3	3.5	4	3
Common Ragweed	4	3	2.5	3
Giant Ragweed	2.5	3	2	3
Cocklebur	3	3	2.5	3

Table 3. Effectiveness ratings of post-emergent herbicides on a series of corn-threatening weeds (1=poor, 2=fair, 3=good, 4=excellent) (Owen, 2013; Ackerman et al., 2014).

	Atrazine	Foramsulfuron + iodoflufenuron	Mesotrione	Flumesulam + clopyralid	Imazethapyr
Foxtail	2	3.5	1	1	2.5
Pigweeds and Waterhemp	4	3	4	3.5	2.5
Lambsquarters	4	3	3	2	3
Velvetleaf	4	3.5	4	3.5	3.5
Common Ragweed	4	4	2	4	3
Giant Ragweed	3	3	3	3.5	2
Cocklebur	4	4	3.5	4	3.5

Increasing reliance on newer herbicides comes with risks, since the full range of environmental consequences associated with their use remains unknown. Moreover, weeds can rapidly develop resistance to novel treatments, even those that are hardly applied. In one experiment conducted at Kansas State University, a patch of Palmer amaranth grew resistant to six separate herbicides, three of which had almost never been used on the plot in the previous four decades (Brown, 2021). Entering a cycle in which scientists must produce new herbicides every few years to combat the increasing resistance of weeds would be both impractical and, from a financial standpoint, unfavorable.

The widespread implementation of integrated weed management

(IWM), a set of low-to-zero chemical strategies designed to combat weeds, is critical in creating a sustainable future for agriculture. IWM practices, which include crop rotation, the use of cover crops, enhanced tillage methods, and intercropping, reduce chemical costs and the environmental consequences of applying pesticides while considerably boosting crop yields. For example, an Iowa study found that corn yields were almost 4% higher in fields that implemented ridge tillage, an IWM method by which crops are planted on elevated mounds while plant residue and weeds are left in shallow rows, in comparison to those using conventional tillage methods (Ackerman et al., 2014).

4. Conclusion

Atrazine exposure is associated with a wide range of adverse health effects, including thyroid and ovarian cancers in applicators, small-for-gestational-age births, premature births, low birth weights, and spontaneous abortions in pregnant women, menstrual irregularity and diminished hormone levels in other women, and reduced sperm motility and quality in men. Recall that the women in both the Ohio singleton birth study and the Illinois and Vermont analysis had been exposed to atrazine concentrations well below the 3 µg/L limit currently needed to enter the EPA's AMP. Regardless, atrazine exposure was significantly associated with damages in those women and, in the case of the Illinois and Vermont study, their newborn children. Research on both laboratory and wild animals, including Hayes et al. (2003), in which frogs displayed hermaphroditism after being exposed to just 0.1 atrazine p.p.b., has consistently found physical harm at doses far lower than the current EPA limit.

Thus, there is evidence to suggest that current atrazine restrictions in the United States are far too lenient. As shown by the CEEPES model, implementing stronger atrazine restrictions near important bodies of water and/or reducing its application to post-emergent situations would greatly reduce usage while only marginally affecting crop yields and prices. The United States should also consider introducing a complete ban on atrazine. After Italy and Germany barred the substance in 1991, no significant changes were seen in the crop yields and production, both of which continued to rise for at least another decade, in both countries. Again, such a policy would have little impact on the economic side of the agricultural industry.

The question of whether the need to protect the health of applicators and the general population from the effects of atrazine outweighs the slight changes in crop productivity that would follow more substantial legislation is subjective, but, given the evidence, I believe we have no reason to avoid stronger environmental policies. We must also note that destructive target weeds are demonstrating expanding resistance to atrazine, a trend that has held true with numerous other pesticides. Waiting until entire maize and sorghum fields begin dying due to the widespread growth of atrazine-resistant "superweeds" would not be economically or ethically ideal. I recommend that growers diversify their arsenal of pesticide treatments in the short-term while working to implement IWM strategies on a larger scale over time.

5. Bibliography

1. Abell A, Ernst E, Bonde JP (2000). Semen quality and sexual hormones in greenhouse workers. *Scand J Work Environ Health*. 2000 Dec;26(6):492-500. doi: 10.5271/sjweh.573. PMID: 11201396.
2. Ackerman, F. (2007). The economics of atrazine. *International Journal of Occupational and Environmental Health*, 13(4), 437-445. <https://doi.org/10.1179/oeht.2007.13.4.437>
3. Ackerman, F., Whited, M., Knight, P. (2014). Would banning atrazine benefit farmers? *International Journal of Occupational and Environmental Health*, 20(1), 61-70. DOI 10.1179/2049396713Y. 0000000054
4. Almberg KS, Turyk ME, Jones RM, Rankin K, Freels S, Stayner LT (2018). Atrazine Contamination of Drinking Water and Adverse Birth Outcomes in Community Water Systems with Elevated Atrazine in Ohio, 2006-2008. *Int J Environ Res Public Health*. 2018;15(9). doi: 10.3390/ijerph15091889
5. Aviv, Rachel (2014). A Valuable Reputation. Retrieved from newyorker.com/magazine/2014/02/10/a-valuable-reputation.
6. Ayotte, P., Giroux, S., Dewailly, É., Avila, M. H., Farias, P., Danis, R., & Díaz, C. V. (2001). DDT Spraying for Malaria Control and Reproductive Function in Mexican Men. *Epidemiology*, 12(3), 366–367. <http://www.jstor.org/stable/3703719>
7. Bostofte, E., Serup, J., & Rebbe, H. (1982). Relation between morphologically abnormal spermatozoa and pregnancies obtained during a twenty-year follow-up period. *International Journal of Andrology*, 5(4), 379.
8. Brown, C. (2021). Attack of the Superweeds. Retrieved from nytimes.com/2021/08/18/magazine/superweeds-monsanto.html
9. Carlsen, E., Giwercman, A., Keiding, N., & Skakkebaek, N. E. (1992). Evidence for decreasing quality of semen during past 50 years. *British Medical Journal*, 305(6854), 609-613. <https://doi.org/10.1136/bmj.305.6854.609>
10. Cole, Allison L., "Co-Removal of Atrazine and Nitrate from Groundwater Using a Mulch Biofilm" (2012). Civil Engineering Theses, Dissertations, and Student Research. 49. <https://digitalcommons.unl.edu/civilengdiss/49>
11. Cragin, L. A., Kesner, J. S., Bachand, A. M., Barr, D. B., Meadows, J. W., Krieg, E. F., & Reif, J. S. (2011). Menstrual cycle characteristics and reproductive hormone levels in women exposed to atrazine in drinking water. *Environmental Research*, 111(8), 1293-1301. <https://doi.org/10.1016/j.envres.2011.09.009>
12. De Roos, A. J., Zahm, S. H., Cantor, K. P., Weisenburger, D. D., Holmes, F. F., Burmeister, L. F., & Blair, A. (2003). Integrative assessment of multiple pesticides as risk factors for non-Hodgkin's lymphoma among men. *Occupational and Environmental Medicine (London, England)*, 60(9), e11-11. <https://doi.org/10.1136/oem.60.9.e11>
13. Donna, A., Crosignani, P., Robutti, F., Betta, P., Bocca, R., Mariani, N., Ferrario, F., Fissi, R., & Bernino, F. (1989). Triazine herbicides and ovarian epithelial neoplasms. *Scandinavian Journal of Work, Environment & Health*, 15(1), 47-53. <https://doi.org/10.5271/sjweh.1882>
14. Dosemeci, M., Alavanja, M. C. R., Rowland, A. S., Mage, D., Zahm, S. H., Rothman, N., Lubin, J. H Hoppin, J. A., Sandler, D. P. & Blair, A. (2002). A

- quantitative approach for estimating exposure to pesticides in the agricultural health study. *The Annals of Occupational Hygiene*, 46(2), 245-260. <https://doi.org/10.1093/annhyg/mef011>
15. European Environment Agency (2019). Climate Change Threatens Future of Farming in Europe. eea.europa.eu/highlights/climate-change-threatens-future-of.
 16. FAOSTAT. (2021). Retrieved from fao.org/faostat/en/#home.
 17. Farr, S. L., Cooper, G. S., Cai, J., Savitz, D. A., & Sandler, D. P. (2004). Pesticide use and menstrual cycle characteristics among premenopausal women in the agricultural health study. *American Journal of Epidemiology*, 160(12), 1194-1204. <https://doi.org/10.1093/aje/kwi006>
 18. Freeman, L. E. B., Rusiecki, J. A., Hoppin, J. A., Lubin, J. H., Koutros, S., Andreotti, G., Zahm, S. H., Hines, C. J., Coble, J. B., Barone-Adesi, F., Sloan, J., Sandler, D. P., Blair, A., & Alavanja, M. C. R. (2011). Atrazine and cancer incidence among pesticide applicators in the agricultural health study (1994-2007). *Environmental Health Perspectives*, 119(9), 1253-1259. <https://doi.org/10.1289/ehp.1103561>
 19. Friend, A. D., Valentini, R., Aubinet, M., Matteucci, G., Heinesch, B., Ciais, P., Allard, V., Soussana, J. F., Manca, G., Schulze, E. D., Viovy, N., Loustau, D., Ourcival, J. M., Papale, D., Bernhofer, C., Knohl, A., Buchmann, N., Grünwald, T., Krinner, G., . . . Sanz, M. J. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 437(7058), 529-533. <https://doi.org/10.1038/nature03972>
 20. Hase, Y., Tatsuno, M., Nishi, T., Kataoka, K., Kabe, Y., Yamaguchi, Y., Ozawa, N., Natori, M., Handa, H., & Watanabe, H. (2008). Atrazine binds to F1F0-ATP synthase and inhibits mitochondrial function in sperm. *Biochemical and Biophysical Research Communications*, 366(1), 66-72. <https://doi.org/10.1016/j.bbrc.2007.11.107>
 21. Hayes, T., Haston, K., Tsui, M., Hoang, A., Haeffele, C., & Vonk, A. (2003). Atrazine-induced hermaphroditism at 0.1 ppb in american leopard frogs (*rana pipiens*): Laboratory and field evidence. *Environmental Health Perspectives*, 111(4), 568-575. <https://doi.org/10.1289/ehp.5932>
 22. Hites, R. A., & Raff, J. D. (2013). *Elements of Environmental Chemistry*. John Wiley & Sons.
 23. Krämer, W., Jeschke, P., Schirmer, U., & Witschel, M. (2013). *Modern Methods in Crop Protection Research*. John Wiley & Sons.
 24. Little, A. B. (1992). Executive Summary of Safety and Toxicity Information - Atrazine. *National Toxicology Program*. Retrieved from https://ntp.niehs.nih.gov/ntp/htdocs/chem_background/exsumpdf/atrazine_508.pdf
 25. MacLennan, P., Delzell, E., Sathiakumar, N., Myers, S., Cheng, H., Grizzle, W., Chen, V., & Wu, X. (2002). Cancer incidence among triazine herbicide manufacturing workers. *Journal of Occupational and Environmental Medicine*, 44(11), 1048-1058. <https://doi.org/10.1097/00043764-200211000-00011>
 26. Martin, S. A., Harlow, S. D., Sowers, M. F., Longnecker, M. P., Garabrant, D., Shore, D. L., & Sandler, D. P. (2002). DDT metabolite and androgens in african-american farmers. *Epidemiology (Cambridge, Mass.)*, 13(4), 454-458. <https://doi.org/10.1097/00001648-200207000-00014>

27. Munger, R. (University of Iowa, Iowa City.), Isacson, P., Hu, S., Burns, T., Hanson, J., Lynch, C. F., Cherryholmes, K., Van Dorpe, P., & Hausler, W. J. (1997). Intrauterine growth retardation in iowa communities with herbicide-contaminated drinking water supplies. *Environmental Health Perspectives*, 105(3), 308-314. <https://doi.org/10.1289/ehp.97105308>
28. Agricultural Health Study (AHS) - Epidemiology Branch. (2019). *National Institute of Environmental Health Sciences*. Retrieved from niehs.nih.gov/research/atniehs/labs/epi/studies/ahs/index.cfm
29. Nikkanen, V. (1979). The effects of vasectomy on viscosity, pH and volume of semen in man. *Andrologia*, 11(2), 123.
30. Owen MDK (2011). The importance of atrazine in the integrated management of herbicide-resistant weeds. Ames, IA, USA: Iowa State Univ.
31. Owen, Micheal D. K. (2013). 2013 Herbicide Guide for Iowa Corn and Soybean Production. *Agriculture and Environment Extension Publications*. 181. https://lib.dr.iastate.edu/extension_ag_pubs/181
32. T. Prato and B.-J. Woo (2008). Evaluation of weed control practices for profitability and reduction of atrazine concentrations in surface runoff. *Journal of Soil and Water Conservation*, 63 (2) 99-104; DOI: <https://doi.org/10.2489/jswc.63.2.99>
33. Ribaudo, Marc O., and Aziz Bouzاهر. "Atrazine: Environmental Characteristics and Economics of Management." *USDA*, 1994, ers.usda.gov/webdocs/publications/40593/33085_aer699.pdf?v=0.
34. Ribaudo, M. O., & Hurley, T. M. (1997). Economic and environmental effects associated with reducing the use of atrazine: An example of cross-disciplinary research. *Journal of Agricultural and Applied Economics*, 29(1), 87-97. <https://doi.org/10.1017/S1074070800007574>
35. Rusiecki, J. A., De Roos, A., Lee, W. J., Dosemeci, M., Lubin, J. H., Hoppin, J. A., Blair, A., & Alavanja, M. C. R. (2004). Cancer incidence among pesticide applicators exposed to atrazine in the agricultural health study. *JNCI : Journal of the National Cancer Institute*, 96(18), 1375-1382. <https://doi.org/10.1093/jnci/djh264>
36. Sengupta, P., & Banerjee, R. (2014). Environmental toxins: Alarming impacts of pesticides on male fertility. *Human & Experimental Toxicology*, 33(10), 1017-1039. <https://doi.org/10.1177/0960327113515504>
37. Singh, S., Kumar, V., Chauhan, A., Datta, S., Wani, A. B., Singh, N., & Singh, J. (2018). Toxicity, degradation and analysis of the herbicide atrazine. *Environmental Chemistry Letters*, 16(1), 211-237. <https://doi.org/10.1007/s10311-017-0665-8>
38. Swanton, C. J., Gulden, R. H., & Chandler, K. (2007). A rationale for atrazine stewardship in corn. *Weed Science*, 55(1), 75-81. <https://doi.org/10.1614/WS-06-104.1>
39. Villanueva, C. M., Durand, G., Coutté, M., Chevrier, C., & Cordier, S. (2005). Atrazine in municipal drinking water and risk of low birth weight, preterm delivery, and small-for-gestational-age status. *Occupational and Environmental Medicine (London, England)*, 62(6), 400-405. <https://doi.org/10.1136/oem.2004.016469>
40. Jenna K. Waggoner, Greg J. Kullman, Paul K. Henneberger, David M. Umbach, Aaron Blair, Michael C. R. Alavanja, Freya Kamel, Charles F. Lynch, Charles Knott, Stephanie J. London, Cynthia J. Hines, Kent W. Thomas, Dale

P. Sandler, Jay H. Lubin, Laura E. Beane Freeman, Jane A. Hoppin, Mortality in the Agricultural Health Study, 1993–2007, *American Journal of Epidemiology*, Volume 173, Issue 1, 1 January 2011, Pages 71–83. <https://doi.org/10.1093/aje/kwq323>

41. Wilcox, A. J. (2010). *Fertility and pregnancy: An epidemiologic perspective* (1st ed.). Oxford University Press.



Mutability and Function of SARS-CoV-2 ORF9b

Xiaoyi (Charlotte) Shi

Author Background: *Xiaoyi Shi grew up in China and currently attends Shanghai Foreign Language School Affiliated to SISU in Shanghai, China. Her Pioneer research concentration was in the field of biology and titled “Pandemic: The New Coronavirus.”*

Abstract

SARS-CoV-2 is the causative agent of the highly transmissible disease COVID-19. Its accessory genes have been suggested as contributors to viral pathogenicity, so understanding the mutability and function of these accessory genes is of great importance for drug design. The accessory gene ORF9b is an open reading frame inside the SARS-CoV-2 nucleocapsid N gene and is translated by the leaky scanning mechanism, which produces a frameshift protein that has been proven to disturb multiple important proteins in the innate immune pathway, including the mitochondria translocase TOM70 and the nuclear factor-kappa B essential modulator NEMO. In this study, the mutability and function of the accessory protein ORF9b were investigated. First, the evolutionary relationship between SARS-CoV-2 ORF9b and other N genes' internal open reading frame proteins belonging to the *Betacoronavirus* genus lineage A and lineage B was analyzed. Then, a software workflow was developed to isolate all the ORF9b variants available via the National Center for Biotechnology Information (NCBI) and the mutability of ORF9b was assessed. Additionally, several potentially beneficial mutations that may increase in frequency as evolution proceeds were predicted based on observed mutation frequency, protein docking studies, and protein stability prediction. Moreover, based on protein docking results, it was postulated that Hsp90 folds the β -sheet-rich ORF9b homodimer into a helical structure so that ORF9b can bind TOM70. Finally, three sets of experiments were proposed to gain deeper insight into the mode of regulation of ORF9b leaky scanning, Hsp90's chaperonic role in ORF9b folding, and the mode of ORF9b-NEMO interaction, which lays the foundation for future investigations of ORF9b on both the RNA level and the protein level. In summary, these findings are crucial stepping stones for comprehending ORF9b's contribution to SARS-CoV-2's pathogenicity and anticipating the evolutionary path that ORF9b may take in the future.

1. Abbreviations

ACE, atomic contact energy
BAX, Bcl-2-associated X
HBV, hepatitis B virus
HCV, hepatitis C virus
HSP, heat-shock protein
IFN, interferon
IPTG, isopropyl β -D-1-thiogalactopyranoside
IRF, interferon regulatory factor
ITC, isothermal titration calorimetry
MAVS, mitochondrial antiviral protein
MCV, molluscum contagiosum virus
NCBI, National Center for Biotechnology Information
NEMO, nuclear factor-kappa B essential modulator
NSP, nonstructural proteins
ORF, open reading frame
ORFV, orf virus
PAMP, pathogen-associated molecular pattern
PINK1, PTEN-induced kinase 1
RIG, retinoic acid-inducible gene
RLR, RIG-I-like receptor
TBK, TANK-binding kinase 1
UTR, untranslated region

2. Introduction

SARS-CoV-2 is the agent of COVID-19, the highly transmissible disease that has spread throughout the globe since it was first reported in Wuhan, China, in December 2019 (WHO, 2021). There have been more than 190 million confirmed cases of COVID-19, including more than four million deaths, reported to the WHO as of July 2021 (WHO, 2021). Since the identification of SARS-CoV-2, numerous efforts have been made to understand the virus in order to meet the global demand for vaccines and medicines. One of these efforts involves elucidating the viral genome. The first SARS-CoV-2 genome (GenBank: MN908947.3) was sequenced and released on January 10, 2020 (Wu, 2020). Since then, more sequences have been released from the National Center of Biotechnology Information (NCBI), Global Initiative on Sharing Avian Influenza Data (GISAID), and European Nucleotide Archive (ENA). Research currently underway has uncovered many features of the SARS-CoV-2 genome. SARS-CoV-2 has a large (+) RNA genome of 30kb flanked by 5' and 3' untranslated regions (UTRs). At the 5' end, ORF1a and ORF1b account for two-thirds of the genome, and the two open reading frames (ORFs) code for 16 nonstructural proteins (NSPs) responsible for replication and the maintenance of the large genome. At the 3' end, structural protein-coding ORFs (M, E, S, and N proteins) and accessory genes account for the remaining one-third of the genome (V'kovski et al., 2020). A comparison between the SARS-CoV-2 and Bat-CoV genomes reveals

that RaTG13 shares the closest evolutionary relationship with SARS-CoV-2 (Boni et al., 2020).

Although accessory genes are not essential for the *in vitro* viral replication cycle and are sometimes considered to be “luxury” genes (Schaecher & Pekosz, 2009), they often provide the virus with some survival advantages in the host environment, contributing to the virulence and pathogenicity of the virus (Lewis et al., 2017). Strong protein-coding signatures have been found for nine accessory protein ORFs in the SARS-CoV-2 genome, including ORF3a, 3b, 6, 7a, 7b, 8, 9b, 9c, and 10 (Jungreis et al., 2020). ORF9b is an open reading frame inside the nucleocapsid (N) gene. It is probably translated through the ribosomal leaky scanning mechanism, where the scanning ribosome bypasses the initial start codon and initiates the translation at a downstream AUG (Xu et al., 2009). ORF9b has proteomics support (Bojkova et al., 2020; Davidson et al., 2020; Bezstarosti et al., 2020), alternate-frame translation support from ribosome footprinting (Finkel et al., 2020), and evidence of cellular protein interaction with ORF9b RNA (Schmidt et al., 2020). The SARS-CoV-2 proteome microarray shows that COVID-19 patients have a strong IgG and IgM antibody response that specifically targets ORF9b (Jiang et al., 2020), suggesting that ORF9b might be highly foreign to humans and therefore highly antigenic. SARS-CoV-2 ORF9b has attracted a lot of interest due to its interference with the retinoic acid-inducible gene (RIG)-I-like receptors (RLR) signaling pathways (Fig. 1).

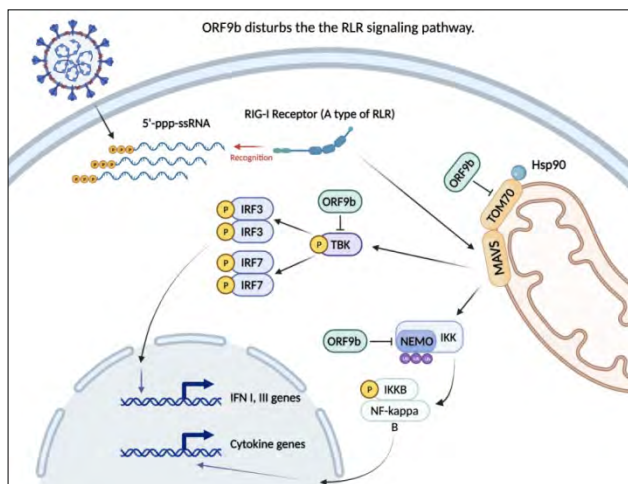


Figure 1. This figure summarizes the RIG-I-like Receptor (RLR) signaling pathway. In this pathway, SARS-CoV-2 ORF9b suppresses the phosphorylation of the protein kinase TBK1, interacts with the mitochondria translocase protein TOM70, and interrupts the K63-linked poly-ubiquitination of the nuclear factor-kappa B essential modulator (NEMO) essential for NEMO activation. ORF9b's interaction with TBK1 inhibits the phosphorylation of interferon regulatory factors IRF3 and IRF7, thereby suppressing the downstream transcription of interferon IFN I, III genes. Its interaction with NEMO inhibits the activation of the downstream kinase IKKB and nuclear factor NF-kappa B activation, thereby suppressing the transcription of cytokine genes.

The RLRs are pathogen-associated molecular pattern (PAMP)-recognizing receptors that are responsible for detecting viruses (Offermanns & Rosenthal, 2008). The RIG-I receptor is one type of RLR involved in viral 5'-triphosphate-RNA recognition based on the presence of 5'-triphosphate as a potent marker of non-self (Loo & Gale, 2011). This is because although RNA polymerase generates 5'-triphosphate for both viral and eukaryotic RNA molecules, eukaryotic RNAs undergo further post-transcriptional modifications including cleavage and capping, which distinguishes between viral RNA and cellular RNA and thereby enables viral RNA recognition based on 5'-triphosphate detection. The binding of the viral 5'-triphosphate with the cell's RIG-I receptor C-terminus induces conformational changes at the RIG-I receptor N-terminus, turning it into an "open" conformation that enables protein-protein interaction. The RIG-I receptor N-terminus contains an interaction motif found in many proteins involved in inflammation, thereby facilitating its downstream interaction with the mitochondrial antiviral protein (MAVS). This pathway eventually activates interferon regulatory factors IRF3 and IRF7 and the transcription factor NK-kappa B to turn on the transcription of Type I and Type III interferons (IFN) (Loo and Gale, 2011).

Recent studies have demonstrated that SARS-CoV-2 ORF9b can interact with multiple proteins at several sites in the RLR pathway, disturbing IFN-1 and IFN-3 production (Jiang & Li, 2020).

(1) ORF9b suppresses the phosphorylation of the protein kinase TBK1. TBK1 is downstream of MAVS and upstream of interferon regulatory factor 3 (IRF3), whose phosphorylation is crucial for IRF3 activation. Consequently, the suppression of TBK1 phosphorylation results in a reduction in IRF3 phosphorylation, thus suppressing its activity as a transcription factor that turns on the interferon gene (Han et al., 2020).

(2) ORF9b interacts with the mitochondria translocase protein TOM70. Besides the translocation of pre-proteins into the mitochondria, TOM70 is also the key adaptor protein in the RLR signaling pathway. Protein crystallography studies demonstrate that ORF9b inhibits TOM70 activity by interacting with its C-terminal domain and thereby allosterically inhibiting its association with HSP90 (a chaperone protein and an important adaptor protein in the RLR pathway) (Gao et al., 2021). It is suggested that ORF9b localizes to the mitochondria and interacts with TOM70, eventually suppressing IFN-1 production (Jiang & Li, 2020).

(3) ORF9b interrupts the K63-linked poly-ubiquitination of nuclear factor-kappa B essential modulator (NEMO). K63-linked poly-ubiquitination is the ubiquitination of lysine at position 63 in the ubiquitin molecule, a signal protein that is conjugated to other intracellular proteins. Although K48-linked polyubiquitination is associated with protein degradation, K63-linked polyubiquitination is associated with other functions such as inflammation (Miranda & Sorkin, 2007). NEMO is a modulator component of the I-kappa kinase (IKK) complex. The K63-linked poly-ubiquitination of NEMO is essential for IKK activation and the phosphorylation of downstream kinase IKKB. This pathway eventually facilitates the translocation of the nuclear factor NF-kappa B into the nucleus as a transcription factor that turns on its target genes (Wu et al., 2021). Experiments on the deletion of secondary structure elements have suggested that one alpha-helix ($\alpha 1$) and two beta sheets ($\beta 1$, $\beta 2$) of ORF9b interact

with NEMO (Wu et al., 2021). It is postulated that the recruitment of NEMO to the mitochondria MAVS complex spatially increases the chance of ORF9b-NEMO interaction (Wu et al., 2021).

It is also intriguing that although SARS-CoV ORF9b and SARS-CoV-2 ORF9b can interact with TOM70 with a similar strength, suggesting the same mode of interaction (Jiang et al., 2020), SARS-CoV ORF9b suppresses the RLR pathway by promoting MAVS/TRAF3/TRAF6 degradation, while SARS-CoV-2 ORF9b targets MAVS-downstream proteins. (Shi et al., 2014). That is to say, SARS-CoV-2 ORF9b and SARS-CoV ORF9b share the strategy of inhibiting TOM70, but they approach other parts of the RLR pathway in different ways. It is also noticeable that ORF9bs from the two viruses share no more than 72% amino acid identity (Wu et al., 2021) but can interact with TOM70 with similar strength.

The complicated interaction between ORF9b and the interferon pathway makes ORF9b a potential drug target. However, the mechanism of these interactions has not been fully elucidated, and the specific motifs and important amino acid residues involved remain uncovered. Understanding the molecular mechanism of the protein-protein interaction that ORF9b involves is crucial for medical applications, and the elucidation of the different strategies that SARS-CoV and SARS-CoV-2 employ to interfere with the interferon pathway is critical for obtaining deeper insights into their respective pathogenicity.

3. Experimental Design

To understand the relationships between ORF9b and other N gene internal ORF proteins, the evolutionary correlation of the N internal ORF proteins across different *Betacoronaviruses* was analyzed, and it was concluded that the strategy of employing an N internal ORF protein to disturb the RLR pathway is unique to the *Betacoronavirus* lineage B.

Then, attention was devoted to mutations in ORF9b and assessing their significance. A software workflow was designed to efficiently isolate all the SARS-CoV-2 ORF9b variants from 150,131 genomes uploaded to the National Center for Biotechnology Information (NCBI). The result revealed that ORF9b is relatively conserved throughout its evolution, and the mutation frequency for every mutation within the residues 43-78 that interacts with TOM70 was calculated. The potentially beneficial mutations in terms of ORF9b-TOM70 interaction were then predicted. Protein docking studies and protein stability trials were carried out to verify the significance of these mutations, and it was predicted that K59N, Q70L, and Q77E might increase in frequency as the evolution proceeds.

To obtain deeper insights into the protein-protein interactions that ORF9b involves, Hsp90's role in ORF9b-TOM70 binding was assessed. It was proposed that Hsp90 might facilitate the folding of the β -sheet-rich ORF9b homodimer into an α -helical structure, thereby enabling ORF9b to structurally accommodate the TOM70 C-terminus and facilitating ORF9b-TOM70 interaction, which in turn inhibits Hsp90's binding to TOM70.

Finally, to better understand both the RNA-level and the protein-level biology of ORF9b, three experiments were proposed to elucidate the leaky scanning regulation, Hsp90's role in ORF9b folding, and the mode of ORF9b-NEMO interaction.

4. Materials and Methods

4.1 Analysis of *Betacoronavirus* N Internal ORFs

Since SARS-CoV-2 ORF9b is an ORF inside the N gene translated through the leaky scanning mechanism, attention was first devoted to other *Betacoronavirus* N internal ORF proteins to better understand their relationships. The *Betacoronavirus* genus is further divided into 4 subgenera, or lineages: lineage A, B, C, and D, and both SARS-CoV and SARS-CoV-2 fall into lineage B. In *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*, genome organization of *Betacoronavirus* shows that *Betacoronavirus* lineage A and lineage B contain N internal ORFs (King, 2012). To determine whether the two lineages' N internal ORFs share the same evolutionary origin, seven N internal ORF protein sequences from the two lineages were downloaded from Uniprot Protein Bank (Protein Bank ID: P22654-1, Q4VID0-1, A0A0A7V3X8-1, Q0ZJG8-1, C3VPF4, P59636, P0DTD2) (Fig. 3). By blasting the RaTG13 genome (a Bat-CoV and the closest relative to SARS-CoV-2) using the SARS-CoV-2 ORF9b as a query, the RaTG13 ORF9b gene was acquired. The gene was then translated into an amino acid sequence using the Expasy translation tool (Gasteiger, 2003). Protein alignment and a phylogenetic tree were then constructed using the Clustal Omega webserver (Sievers et al., 2011).

4.2 Collection of SARS-CoV-2 ORF9b variants

To collect the SARS-CoV-2 ORF9b variants, complete genomes with no ambiguous characters collected from January 2020 to June 2021 were downloaded from NCBI using the filters "SARS-CoV-2, taxid:2697049," "Ambiguous Characters: Max: 0," "Sequence Type: GenBank," "RefSeq Genome Completeness: complete," and "Nucleotide Completeness: complete." A full-length genome collection with no ambiguous characters was thus established for each month between January 2020 and June 2021. A software workflow was then designed to process the genome information (Fig. 2). The TBtools tool kit was used to process genome information (Chen et al., 2020); the Mega version X program was used for protein alignment (Kumar, Stecher, Li, Knyaz, and Tamura 2018; Stecher, Tamura, and Kumar 2020); the sequence logo was created using WebLogo (Crooks, 2004); and the mutation frequency was calculated in Jalview (Waterhouse et al., 2009).

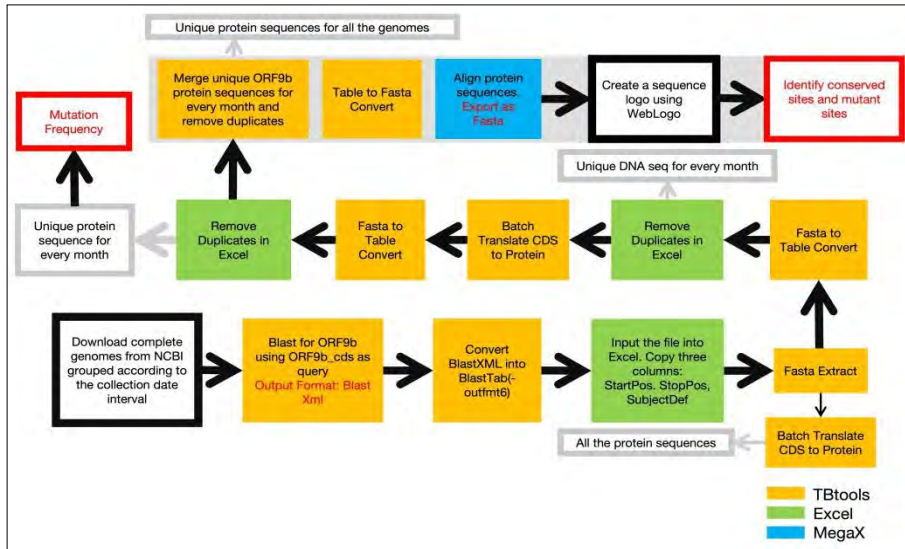


Figure 2. A software workflow was designed to process viral genomes. Boxes colored in yellow, green, and blue involve the use of software TBtools, Excel, and MegaX, respectively. This workflow could eventually yield all the unique ORF9b protein sequences and DNA sequences.

4.3 Predicting the Potentially Beneficial Mutations in terms of TOM70 binding

It was assumed that a mutation that survives in the genome collection or that rises in frequency is likely neutral, beneficial, or slightly deleterious. To identify the mutations that are likely beneficial, mutations that were observed in at least 6 months' collection or those which displayed a significant increase in frequency across time were collected, and attention was focused on the regions involved in protein-protein interaction with TOM70 (amino acid positions 43-78). However, these mutations might be deleterious, since ORF9b is an accessory gene not essential for viral replication; it is also possible that beneficial mutations were lost during evolution because of genetic drift and the founder effect. Additionally, mutation frequency might not be accurate due to statistical bias. To diminish statistical noise and verify the functional significance of these mutations, protein docking studies were carried out using the PatchDock server to evaluate the degree to which ORF9b-TOM70 interaction is favorable for each selected mutation (Schneidman-Duhovny et al., 2005, Duhovny et al., 2002). Then, I-Mutant Suite was used to evaluate the effect of point mutations on protein stability (Capriotti et al., 2005).

4.4 A Closer Look at the ORF9b-TOM70 Interaction

The protein structure of the middle domain and the C-terminus of Hsp90 were downloaded from Protein Data Bank (PDB: 3Q6N). Protein Docking studies were carried out using the PatchDock server (Schneidman-Duhovny et al., 2005;

Duhovny et al., 2002). Docking results were visualized in PyMol (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC. 2011).

5. Results

5.1 ORF9b is Unique to *Betacoronavirus* Lineage B

Five N gene internal ORF proteins from *Betacoronavirus* lineage A and three from lineage B were collected from Uniprot (Fig.3A). The protein alignment indicated that the proteins from the two lineages are very different, both in length and content. The phylogenetic tree revealed that the two lineages separated on two distinct areas on the graph (Fig. 3B).

It can thus be concluded that the strategy of employing an N gene internal ORF protein to disturb the innate immune system RLR pathway is unique to *Betacoronavirus* lineage B, which includes the SARS-CoV and SARS-CoV-2 virus.

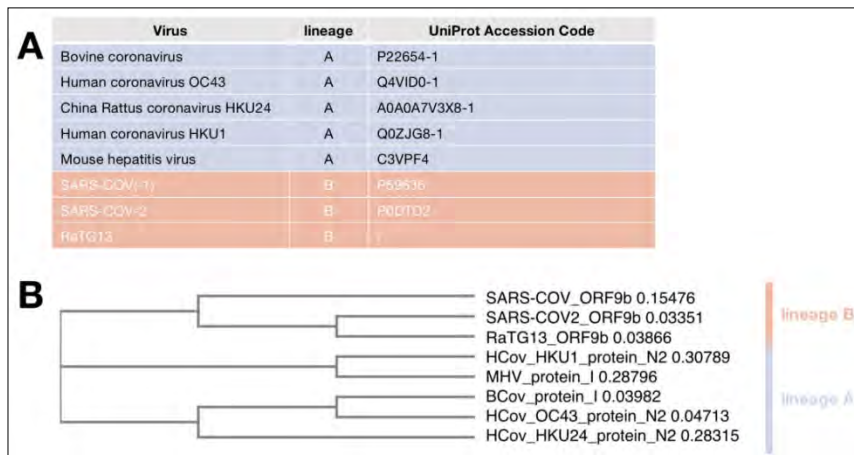


Figure 3. A. The UniProt accession codes of five N internal reading frame proteins from *Betacoronavirus* lineage A and three N internal reading frame proteins from lineage B. **B.** The distribution of the 8 protein sequences on the phylogenetic tree revealed a different evolutionary origin of the N internal open reading frame protein from the two lineages.

5.2 Collection of SARS-CoV-2 ORF9b variants

A total of 150,131 complete genomes were downloaded from the NCBI database, and they were classified by their collection date (Fig. 4a). ORF9b variants were isolated from this large collection and unique ORF9b coding sequences, as well as unique ORF9b proteins were collected according to the software workflow described in Fig. 2.

A sequence logo was created using WebLogo (Fig. 4b), showing that ORF9b is conserved for most residues. On the other hand, several variants were

observed to contain no ORF9b due to a deletion mutation at the ORF9b leaky scanning start codon, indicating that ORF9b is not required for viral survival. Still, ORF9b's ability to disturb the immune system by targeting multiple proteins in the RLR pathway, as well as the high level of ORF9b-specific antibodies in patients' blood, provides support for ORF9b's important role in conferring survival advantages for the virus. Moreover, considering the structural importance of the N gene from which ORF9b is derived, it is possible that the conservation of ORF9b is because of dual factors: the conservation of the important nucleocapsid gene essential for viral survival, and ORF9b itself, which confers a survival advantage to the virus.

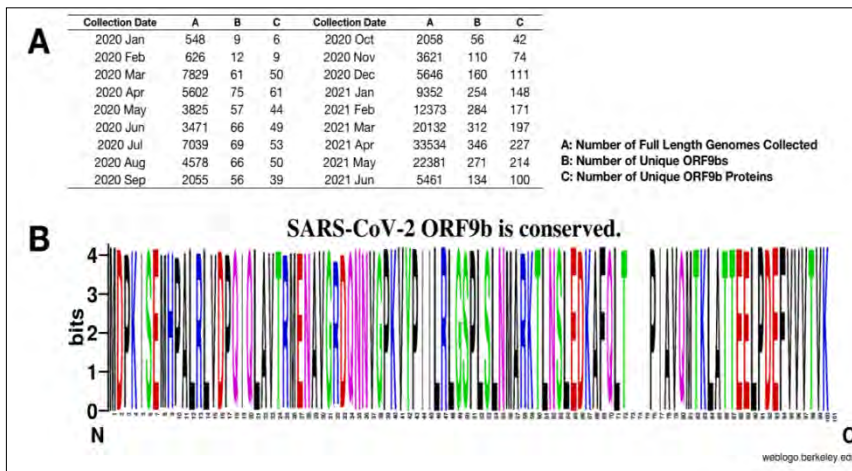


Figure 4. A. The number of full-length SARS-CoV-2 genomes, number of unique ORF9bs, and number of unique ORF9b proteins were recorded and classified according to the collection date interval. **B.** The sequence logo revealed that SARS-CoV-2 ORF9b is conserved.

5.3 K59N, Q70L, and Q77E are Potentially Beneficial Mutations

Attention was first devoted to residues 43-78 which were demonstrated to be bound to the TOM70 protein involved in the RLR signaling pathway (Gao et al., 2021). The mutation frequency was calculated in Jalview and the timeline of evolution was recorded for each residue between 43 and 78 (Supplementary Material). The calculation result revealed that ORF9b is conserved for most sites, and only two mutations exceeded the frequency of 0.5%: T60A and Q77E.

Then, the timeline was more carefully examined, and mutations that appeared in at least six months' collection or those that experienced a significant increase in frequency across time were selected. Twenty-four mutations satisfied this criterion: R47L, P51L, L52S, S53L, M56I, A57V, R58K, K59N, K59N, T60L, T60A, N62S, N62D, L64F, K67R, A68V, Q70L, T72I, P73S, I74*, A75V, Q77H, Q77E, and M78T. Among the 24 mutations, T60A and Q77E exhibited a significant increase in frequency after May 2021, and Q77E displayed nearly exponential growth (Fig. 5).

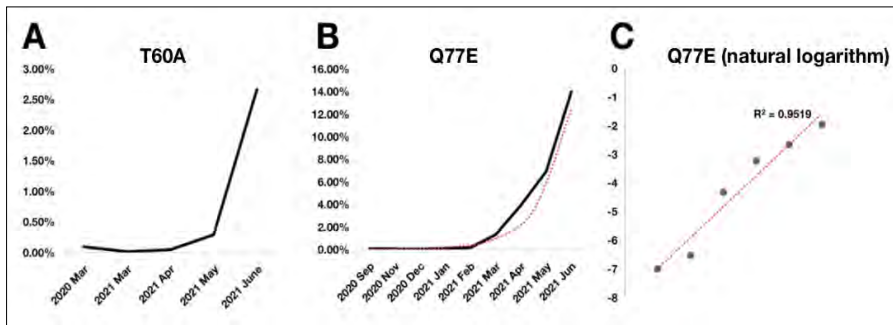


Figure 5. A. T60A first appeared in March 2020 and was not observed again until March 2021. It exhibited a significant increase in frequency since May 2021. **B.** Q77E displayed nearly exponential growth since September 2020 and was observed in the genome collection for consecutive nine months. (The red curve indicates an exponential trend line.) **C.** Scatterplot of the natural logarithm of the Q77E mutation rate. The trend line has a coefficient of determination $R^2=0.9519$, indicating a high degree of linear correlation between the natural logarithm of the Q77E mutation rate and time, revealing exponential growth of this mutation over time.

To understand the functional significance of those mutations in terms of ORF9b-TOM70 interaction, the 24 point mutations identified above were introduced into ORF9b from the reference genome and protein models of the 24 ORF9b variants were then constructed using the crystal structure of SARS-CoV-2 ORF9b complex with human TOM70 (Protein Data Bank (PDB) ID: 7DHG) as a template. Then, protein docking studies were carried out using PatchDock. TOM70 (PDB: 7DHG) was input as the receptor, and the ORF9b variant protein models were input as ligands, respectively. PatchDock provides solutions to protein-protein docking sorted by the geometric shape complementary score, which indicates to what extent the ligand and receptor match geometrically; the higher the score, the better the match. Of the 24 ORF9b variants, 7 variants (P51L, K59N, N62D, Q70L, A75V, Q77H, and Q77E) made geometric shape complementary scores higher than the reference ORF9b (Fig. 6), suggesting better geometric complementarities. Surprisingly, T60A had a much lower score (21432) than ORF9b (22546), suggesting that its significant increase in frequency is either because of genetic drift or survival advantages yet to be determined. Besides the complementary score, approximate interface area and atomic contact energy (ACE) were also acquired, and the latter refers to the energy change upon replacing the water molecules in protein-water interaction with the protein ligand. (A lower ACE signals a more favorable interaction.)

To better understand how these 7 point mutations affect protein stability, the ORF9b monomer from an X-ray crystallographic study (PDB: 6Z4U) was used as a wild type ORF9b template, and $\Delta\Delta G$ upon single point mutation was evaluated to predict whether the mutation is favorable in terms of protein stability. $\Delta\Delta G$ is a metric that measures the change in the Gibbs free energy of folding of a protein upon point mutation:

$$\Delta\Delta G = (G_{unfolded} - G_{folded})_{WT} - (G_{unfolded} - G_{folded})_{Mutant}$$

A positive $\Delta\Delta G$ means stabilization, while a negative $\Delta\Delta G$ means destabilization. The condition was set as 37°C and pH=7.4 to mimic the human cell environment

Mutations	Shape Complementarity Score	Complex Interface Area	Atomic Contact Energy(kcal/mol)	$\Delta\Delta G$ (kcal/mol)
WT	22546	2446.30	-481.28	/
K59N	23996	2668.70	-494.64	+0.16
Q77E	23496	2499.30	-370.21	-0.28
P51L	23252	2607.70	-530.46	-0.52
N62D	23238	2597.00	-526.85	-0.12
Q70L	23178	2611.30	-569.70	+0.38
A75V	23108	2544.20	-531.33	-0.11
Q77H	22640	2670.10	-481.57	-0.75

Figure 6. Protein docking results and protein stability change upon point mutation ($\Delta\Delta G$) were recorded. A positive $\Delta\Delta G$ means stabilization, while a negative $\Delta\Delta G$ means destabilization.

The results from the two trials were then examined in depth. Attention was focused on three variants: K59N, which scored highest in the protein docking studies, Q70L, which scored highest in the protein stability trial, and Q77E, whose frequency exhibited exponential increase over time.

The K59N variant scored the highest (23996) among the 7 variants in the protein docking trial, characterized by a large interface area (2668.70) of atomic contact. It also displayed a lower ACE (-494.64 kcal/mol) compared to the reference ORF9b upon docking, which means lower desolvation energy and a more favorable interaction. Moreover, introducing K59N into the ORF9b monomer yielded a positive $\Delta\Delta G$ (+0.16 kcal/mol), suggesting that it stabilizes the protein structure.

The Q70L variant scored significantly higher (23178) than the wild type ORF9b (22546), and it was predicted to have a large interface area (2611.30) that interacts with TOM70. It displayed a much lower ACE of -569.70 kcal/mol than the wild type (-481.28 kcal/mol) and K59N. More notably, introducing Q70L to the protein yielded the largest $\Delta\Delta G$ (+0.38 kcal/mol), suggesting its strong ability to stabilize the protein monomer.

Q77E scored second among the 7 variants in terms of geometric complementarity (23496), and it scored significantly higher than the reference ORF9b (22546), which accords with its exponential increase in the genome collection over time. Although the mutation yielded a negative $\Delta\Delta G$, suggesting that it destabilizes the protein monomer, and its ACE (-370.21) was higher than the wild type ORF9b upon interaction with TOM70, it is still safe to postulate that Q77E is a beneficial mutation for SARS-CoV-2 because it enhances the chance of interaction with TOM70 by modifying its shape to better accommodate with TOM70.

Considering the potential ability for the K59N and Q70L variants to interact better with TOM70 and their capability to stabilize the protein monomer, it is reasonable to predict that these two mutations might increase over time.

Additionally, Q77 was also predicted to keep increasing in frequency due to its advantage in geometric complementarity with TOM70 and its large percentage in the gene pool.

5.4 Hsp90 May Fold the β -sheet-rich ORF9b into a Helical Structure

ORF9b often exists in the form of β -sheet-rich homodimers in human cells. Surprisingly, it has been demonstrated in previous studies that ORF9b takes up a strikingly different α -helical structure upon binding with TOM70. Moreover, the homodimer ORF9b cannot form a stable structure with TOM70 *in vitro*, and only the helical-structured ORF9b can bind TOM70 and allosterically inhibit Hsp90's binding (Gao et al., 2021). These unexpected phenomena hint that ORF9b must undergo a certain conformational change or folding process to enable its interaction with TOM70.

Hsp90 is a chaperone protein that facilitates protein folding: for example, it helps steroid hormone receptors fold into an active form (Young et al., 2001). Hsp90 has three domains: an N terminal domain responsible for nucleotide-binding, a middle domain for chaperone-client binding, and a C-terminal domain for co-chaperone recognition and homodimerization (Lackie et al., 2017). Both the middle domain and C-terminal domain are involved in protein binding: the middle domain interacts with the protein kinase Akt (Sato et al., 2000), the nitric oxide synthase eNOS (Fontana et al., 2002), and several other proteins, while the C-terminal domain binds proteins such as immunophilins (Young et al., 1998) and TOM70 (Gao et al., 2021). Is Hsp90 responsible for ORF9b folding through the interaction with its middle domain or C-terminal domain? Considering the fact that Hsp90 is recruited to the outer mitochondria membrane upon viral infection (Gao et al., 2021), it would be spatially favorable for Hsp90 to fold ORF9b since the folded α -helical ORF9b can directly interact with TOM70 on the mitochondria.

To examine whether this postulation is reasonable at the molecular level, protein docking studies were performed. The ORF9b homodimer in β -sheet-rich form (PDB: 6Z4U), the ORF9b monomer in β -sheet-rich form (PDB: 6Z4U), and the ORF9b in the α -helical form were input as the ligands, respectively, to dock with the middle domain and C-terminal segment of Hsp90 (PDB: 3Q6N). Geometric complementarity, complex interface area, and atomic contact energy (ACE) were measured. (Geometric complementarity measures the degree to which ORF9b can geometrically accommodate Hsp90; complex interface area often correlates directly with the binding affinity; a lower ACE signals a more favorable interaction.)

ORF9b	Shape Complementarity Score	Complex Interface Area	Atomic Contact Energy(kcal/mol)
β -sheet-rich monomer	15620	2142.70	350.75
β -sheet-rich homodimer	16710	2168.80	-267.25
α -helix	13780	1885.10	-13.73

Figure 7. Protein docking results of different forms of ORF9b upon binding Hsp 90.

Based on these results, the ORF9b homodimer displayed the greatest shape complementarity with Hsp90 compared with the other two forms, and it

had the lowest atomic contact energy at -267.25 kcal/mol, suggesting a favorable interaction. The β -sheet-rich monomer, on the other hand, had a lower shape complementarity with Hsp90, a smaller interface area, and a much higher atomic contact energy, suggesting a much more unfavorable interaction. These results indicated that Hsp90 favors the ORF9b homodimer over the β -sheet-rich ORF9b monomer upon protein interaction. It can thus be postulated that ORF9b in the form of a β -sheet-rich homodimer, instead of its β -sheet-rich monomer, binds Hsp90. Additionally, ORF9b in the form of α -helix had the lowest complementarity with Hsp90, and the atomic contact energy was much higher than ORF9b in the β -sheet-rich homodimer upon binding with Hsp90, suggesting an unstable interaction with Hsp90. Therefore, the helical ORF9b is likely to be released from Hsp90 upon ORF9b-Hsp90 interaction. The favorable interaction between the β -sheet-rich ORF9b homodimer and Hsp90 and the unfavorable interaction between the helical ORF9b and Hsp90 support the postulation that the β -sheet-rich homodimer binds Hsp90 followed by folding and the release of an α -helical ORF9b.

Then, the mode of Hsp90-ORF9b binding predicted in the protein docking studies was examined in the software PyMol (Fig. 8). The β -sheet-rich ORF9b homodimer interacts with the C-terminus of the Hsp90 middle domain (Fig. 8A), while the β -sheet-rich monomer and the α -helix ORF9b tend to bind the N-terminus of the Hsp90 middle domain (Fig. 8B, C). Figure 8D and Figure 8E show the conformational change of ORF9b upon binding with Hsp90.

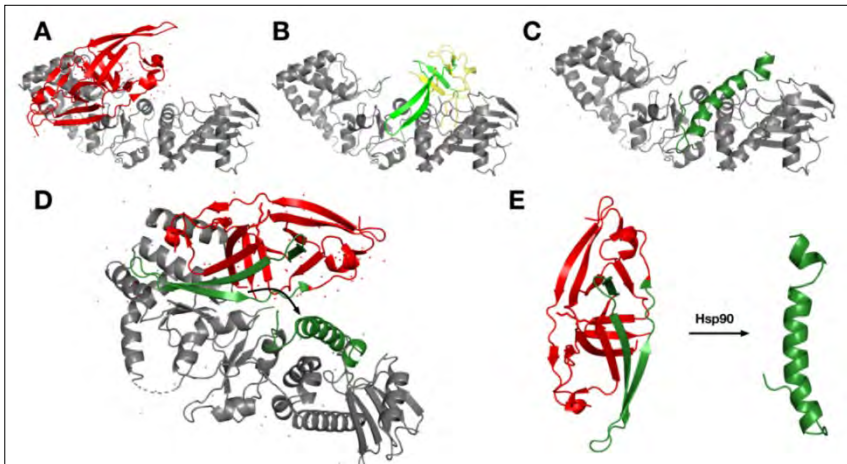


Figure 8. **A.** Homodimeric ORF9b (red) bound to Hsp90 (grey). **B.** β -sheet-rich ORF9b monomer (yellow) bound to Hsp90 (grey). The region colored in green corresponds to the region that interacts with TOM70. **C.** α -helical ORF9b monomer bound to Hsp90. **D.** Superimpose Fig. 8A and Fig. 8C. The region colored in green corresponds to the region that interacts with TOM70. **E.** Fig. 8E illustrates the folding of the region in ORF9b that interacts with TOM70.

Based on the analyses above, a scenario of how ORF9b manages to bind TOM70 was proposed (Fig. 9). (1) Hsp90 folds a part of the ORF9b homodimer into an α -helical structure. (2) The α -helical ORF9b binds to the TOM70 C-terminal domain. (3) Consequently, the interaction between ORF9b and TOM70 allosterically inhibits the binding of Hsp90 to TOM70 (Gao et al., 2021), thereby disturbing the RLR pathway. This postulation speculated that the ORF9b homodimer not only binds TOM70 to interfere with the TOM70-Hsp90 interaction critical in the RLR pathway but also “exploits” Hsp90 to fold it into a helical structure in order to geometrically accommodate TOM70.

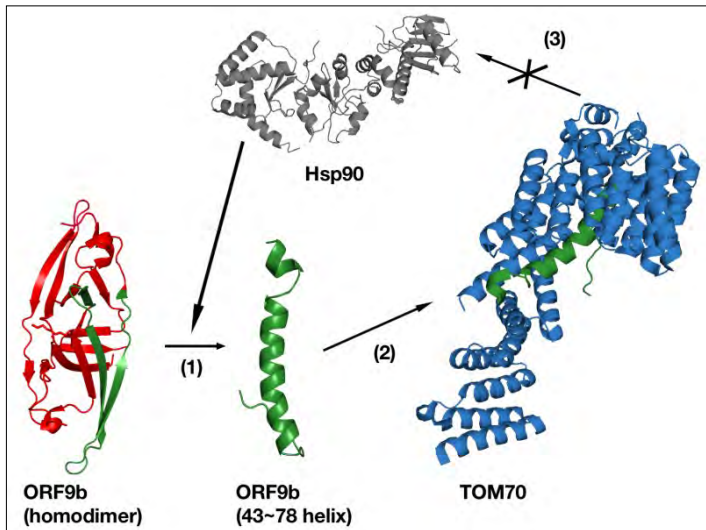


Figure 9. Illustration of the postulated interaction between Hsp90, ORF9b, and TOM70. (1) Homodimeric ORF9b is folded into a helical structure by Hsp90. (2) The helical ORF9b binds TOM70. (3) ORF9b-TOM70 interaction allosterically inhibits Hsp90-TOM70 interaction.

6. Further Research Proposal

To date, some questions regarding the biology of ORF9b on the RNA level as well as the protein level have yet to be answered. One of the questions on the RNA level involves how the leaky scanning mechanism is regulated. On the protein level, despite the protein docking results that support the Hsp90 theory, whether Hsp90, in fact, participates in ORF9b folding has not been demonstrated *in vitro* or *in vivo*. And regarding ORF9b's interruption of the polyubiquitination of NEMO, the detailed mode of interaction has not been elucidated and present knowledge about the mode of interaction is almost absent. Here, several experiments to explore the leaky scanning regulation, the Hsp90 theory, and the mode of ORF9b-NEMO interaction were devised, aiming to offer some ideas for future investigation to fill the vacancy in the biology of ORF9b on both the RNA and protein level.

6.1 Leaky Scanning Regulation

ORF9b is translated through the leaky scanning mechanism (Xu et al., 2009), and its translation initiates 10 nucleotides downstream of the N gene start codon. Kozak consensus sequence is a nucleic acid motif that functions as the eukaryotic translation initiation site. It was discovered that the motif **GCCRCCAUGG** is conserved for vertebrate translation initiation sites (Kozak, 1987). (R refers to a purine.) Start codons (underlined) with a better match to the Kozak consensus sequence are regarded as having an optimal Kozak context or “stronger” start codons and, therefore, are more likely to be translated from. Moreover, the +4 position G and the -3 position R (printed in bold) have been demonstrated to be the most important factor contributing to the strength of the Kozak context. Leaky scanning happens if the upstream start codon provides an inferior Kozak context, making the scanning complexes bypass the initial start codon and start the translation at a downstream start codon with a stronger Kozak context (Stacey et al., 2000). It is the case with the N gene and ORF9b (Fig. 10). The N gene translation initiation site contains a less optimal Kozak context (+4: T; -3:A) compared to ORF9b (+4:G; -3:A) since ORF9b has guanine as its +4 position, which is a better match to the Kozak consensus sequence.

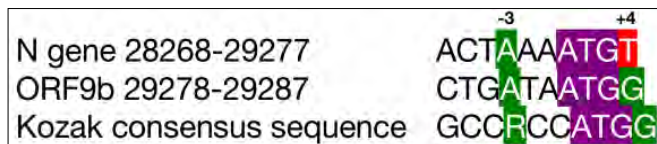


Figure 10. The start codons are colored in purple; favorable Kozak contexts are colored in green, and unfavorable ones are colored in red. The N gene contains -3: A, +4: T, and ORF9b contains -3: A at the translation initiation site. ORF9b provides a more optimal Kozak context for leaky scanning to occur.

It has been demonstrated that the translation level of the downstream ORF (where leaky scanning takes place) is affected by the upstream ORF, as in the case of SARS-CoV ORF9b (Xu et al., 2009). This leads to the question of how SARS-CoV-2 ORF9b translation is regulated and to what extent its level is affected by the Kozak context of the upstream N gene. A set of experiments were proposed to better understand the mode of regulation of ORF9b translation in a Kozak context.

The first step is to transfect HEK 293T cells with plasmids containing the N gene and ORF9b gene mutants described in Fig. 11. The expression level of ORF9b and the nucleocapsid can then be assessed using Western blot analysis. By comparing the expression level of ORF9b and the nucleocapsid for different mutants, the mode of regulation for this leaky scanning can be elucidated.

1	----- CTGATAATGGACCCCA	ORF9b (WT)
2	ACTAAAATGTCTGATAATGGACCCCA	N (WT)
3	ACTAAAATGTCTGATAACGGACCCCA	N (9b-)
4	ACTAAAACGTCTGATAATGGACCCCA	N (N-)
5	ACTAAAATGGCTGATAATGGACCCCA	N (+4T to G)
6	ACTAAACATGTCTGATAATGGACCCCA	N (-3A to C)
7	ACTAAACATGGCTGATAATGGACCCCA	N (-3A to C, +4T to G)
8	ACTAAAATGTCTGATAATGTACCCCA	N (ORF9b +4G to T)
9	ACTAAAATGTCTGATTATGGACCCCA	N (ORF9b -3A to T)
10	ACTAAAATGTCTGATTATGTACCCCA	N (ORF9b -3A to T, +4G to T)

Figure 11. The translation initiation site of the N gene mutants designed for the experiment. Mutants 3~4 disrupt the start codon of ORF9b and the N gene (the deactivated start codon shown in red), respectively. Mutants 5~7 strengthen the N gene Kozak context (mutations shown in purple), and Mutants 8~10 weaken the ORF9b Kozak context (mutations shown in purple).

6.2 The Hsp90 Chaperone Theory

In order to validate Hsp90's chaperonic role in folding ORF9b into a helical structure, a set of experiments were proposed. First, ORF9b, HSP90, and TOM70 will be expressed in *E.coli*. ORF9b will be incubated with and without Hsp90, followed by an isothermal titration calorimetry study to evaluate ORF9b-TOM70 interaction. If the experimental group ORF9b can interact with TOM70, but the control group ORF9b (which is not incubated with Hsp90) cannot, the chaperone role of Hsp90 in ORF9b functional folding can be verified.

(1) Preparing ORF9b, HSP90, and TOM70 for interaction: construct expression plasmids pET28a containing the full-length SARS-CoV-2 ORF9b, the cDNA of TOM70, and a truncated Hsp90 cDNA, coding the middle domain and the C-terminal domain, respectively, following codon optimization. Plasmids will then be transformed into the Escherichia coli protein expression strain BL21 for recombinant protein expression. The *E.coli* transformants will be induced with 0.5 mM Isopropyl β -D-1-thiogalactopyranoside (IPTG) when the cultures grow to OD₆₀₀=0.8. The induction will then be carried out at 18 °C for 18 h. The protein can then be purified as previously described (Gao et al., 2021).

(2) ORF9b folding by Hsp90: Incubate the purified ORF9b with and without Hsp90. Size-exclusion chromatography can then be used to separate Hsp90 and ORF9b in the assay solution based on their differences in size.

(3) ORF9b-TOM70 interaction: Isothermal titration calorimetry (ITC) described in the previous paper can be used to verify whether the ORF9b dimer can bind to TOM70 following Hsp90 folding by measuring the heat released or absorbed upon protein-protein interaction (Gao et al., 2021).

6.3 The Mode of ORF9b-NEMO Interaction

Besides TOM70, ORF9b can also bind NEMO and inhibit its poly-ubiquitination. However, the detailed mode of interaction remains unknown. Still, previous studies on other viruses can perhaps offer some clues. To date, several viral proteins have been demonstrated to be involved in inhibiting the K63-polyubiquitination of NEMO to disturb the RLR pathway, including the

nonstructural protein NS3 of the hepatitis C virus (HCV) (Chen et al., 2015), the core protein HBeAg of the hepatitis B virus (HBV) (Wang et al., 2019), and the MC159 protein of the molluscum contagiosum virus (MCV) (Biswas & Shisler, 2017). Considering the distant evolutionary relationships between MCV (poxvirus), Hepatitis virus (hepadnavirus), and SARS-CoV-2 (coronavirus), their ability to target the K63-polyubiquitination of NEMO is an amazing demonstration of convergent evolution in viruses, indicating that suppressing K63-polyubiquitination of NEMO is crucial for viral pathogenicity. However, previous studies have revealed that these proteins might inhibit the polyubiquitination of NEMO in different ways.

NS3 of HCV inhibits NEMO polyubiquitination by outcompeting NEMO for the ubiquitin chain (Chen et al., 2015). On the other hand, MC159 of MCV inhibits NEMO polyubiquitination by binding the N terminus of NEMO (amino acid position 1-250) (Biswas & Shisler, 2017), thus blocking the cellular inhibitor of apoptosis protein 1's (cIAP1) binding to NEMO at the N terminus. Since cIAP1 is a ligase responsible for ligating the ubiquitin molecules (Jin et al., 2009), disturbing cIAP1-NEMO interaction inhibits the k63-ubiquitination of NEMO. Immunoprecipitation has demonstrated that HBeAg can also bind NEMO, but a NEMO truncate containing only amino acids at positions 1-250 is not sufficient (Wang et al., 2019). It has been indicated that amino acids 51-201 and 306-350 of NEMO are necessary for HBeAg-NEMO binding, suggesting a different mode of interaction with NEMO between HBeAg and MC159 (Wang et al., 2019). Considering the important role that the poly-ubiquitination of NEMO plays in the RLR signaling pathway, elucidating how SARS-CoV-2 ORF9b interacts with NEMO is of great importance for drug design.

Here, a set of experiments that may help determine whether ORF9b takes up a similar strategy as HBeAg or MC159 to interact with NEMO were presented. Notably, the mode of interaction that NS3 employs may not fit into the context of ORF9b, since ORF9b, like HBeAg and MC159, has been demonstrated to directly bind NEMO (Wu et al., 2021), while NS3 binds the ubiquitin chain instead. To determine whether ORF9b adopts the same approach as HBeAg or MC159, truncated NEMO proteins and ORF9b are expressed in *E.coli*, followed by a co-immunoprecipitation study to detect ORF9b-NEMO interaction.

(1) Coexpress NEMOs with ORF9b: Coexpress ORF9b with the human influenza hemagglutinin (HA) tag at its N terminus and FLAG-tagged NEMO truncates as described in Fig. 12 via the expression vector pcDNA3.1.

(2) Co-immunoprecipitation study: transfect the constructs into HEK293T cells, and carry out immunoprecipitation using the method described in Wu et al. (2021).

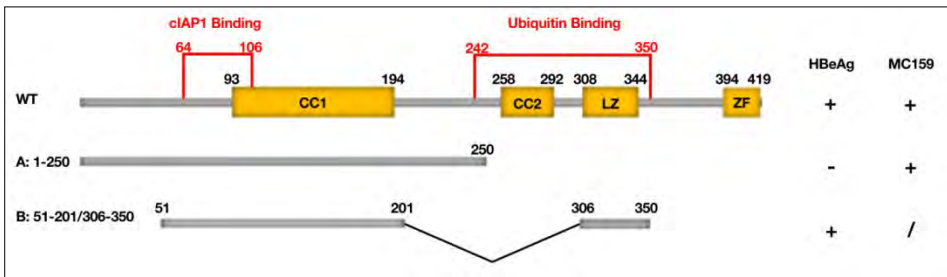


Figure 12. Wild Type (WT) NEMO and its domains. Amino acids 64-106 are involved in cIAP1 binding, and 242-350 are involved in ubiquitin binding. NEMO truncate A (1-250) and NEMO truncate B (51-201/306-350) were designed to test whether ORF9b adopts the same strategy as HBeAg and MC159 to inhibit NEMO polyubiquitination. Previous studies have demonstrated that HBeAg can interact with truncate B but not truncate A, and MC159 can interact with truncate A.

If a detectable level of NEMO truncate A but not B is immunoblotted in the immunoprecipitation products of ORF9b, then ORF9b might be able to interact with NEMO in a way similar to that of HBeAg. On the other hand, if NEMO truncate B but not A is immunoblotted in the ORF9b immunoprecipitation product, it is reasonable to postulate that ORF9b may inhibit k63 poly-ubiquitination of NEMO by disturbing cIAP1-NEMO interaction. To further verify this postulation, a luciferase reporter assay can be carried out to quantify NF-kappa B activation. Since previous reports reveal that overexpression of cIAP1 activates the NF-kappa B pathway (Tang et al., 2003), if the transfected HEK 293T cells containing ORF9b display much lower NF-kappa B luciferase activity compared to the vector control, and the NF-kappa B luciferase activity remains at a constant low level when different doses of cIAP1 are transfected into the cell (while the control groups show increasing luciferase activity as cIAP1 doses increase), the postulation that ORF9b disturbs cIAP1-NEMO interaction stands.

7. Discussion

In this study, the evolutionary relationships between ORF9b and other N internal ORF proteins were first assessed. The N internal open reading frames of *Betacoronavirus* lineage A and lineage B were compared, leading to the conclusion that only *Betacoronavirus* lineage B employs the N internal open reading frame ORF9b protein to evade the immune system. Then, attention was focused on ORF9b mutants and how these mutations affect protein function. A software workflow was designed to easily isolate the ORF9b variants from the complete genome collection and determined that ORF9b was conserved due to dual factors—the importance of the N gene and the survival advantage that ORF9b confers to the virus. Twenty-four potentially beneficial mutations were collected based on their mutation frequency, and Q77E showed an exponential increase over time. The significance of these mutations was then assessed by evaluating their ability to bind TOM70, 7 of them scoring higher than the reference genome ORF9b in terms of geometric complementarity with TOM70. Notably, K59N scored the highest among the 7 mutants. Further analyses of the

change in protein stability upon point mutation revealed that Q70L can stabilize the protein monomer to the greatest extent. Combined data on mutation frequency, protein-interaction ability, and protein stability led to the prediction that K59N, Q70L, and Q77E might increase in frequency as evolution proceeds. To better understand how ORF9b-TOM70 interaction occurs, it was postulated based on protein docking studies that SARS-CoV-2 ORF9b not only manipulates TOM70 to allosterically inhibit Hsp90 binding but also employs Hsp90 to fold itself into a helical structure to accommodate the C-terminal domain of TOM70. Finally, experiments were designed to elucidate the mode of leaky scanning regulation, verify Hsp90's role in ORF9b folding, and explain the machinery of interaction between NEMO and ORF9b, paving the way for future investigations of both the RNA level and the protein level biology of ORF9b.

Admittedly, there are some methodological limitations to this study. First, although a large genome collection consisting of 150,131 full-length genomes was utilized, statistical bias upon genome collection might still have impacted the quality and stringency of the results. For example, some mutations in ORF9b that are beneficial for viral survival in terms of disturbing the immune pathway might have been left out. Moreover, the analyses in this study were based on protein models, so crystallography studies should be carried out in future research to fully verify these conclusions. Additionally, ORF9b monomer stability upon point mutation was evaluated in this study due to the restrictions in the current technique of protein stability calculation; however, ORF9b usually exists in the form of homodimers in human cells, so the stability change calculated in this study should be interpreted carefully.

Previous studies on the evolutionary dynamics of viral proteins involve complex workflows using many software programs to isolate protein variants from large collections of viral genomes. Here, a repeatable software workflow that is easy to manipulate was designed, which may facilitate future research on the evolution of other viral proteins. This study also provides insights into the biology of ORF9b. Since ORF9b is a potential drug target due to its ability to disturb the innate immune pathway at many points, it is critical to understand the biology of this protein. Evaluating the mutability of ORF9b and the significance of these mutations on protein functions is a key step in understanding its biology in depth. With the increasing number of SARS-CoV-2 variants, an omnipotent vaccine to immunize humans from all the variants is almost impossible. Humans have to coexist with SARS-CoV-2 and living with it is a matter of variants and evolution. Therefore, it is especially important to anticipate the evolutionary path SARS-CoV-2 will take. Here, several mutations that might increase in frequency as time goes by were predicted, which provides information for reference to anticipate SARS-CoV-2's evolution in the future.

It has been reported that a dynamic protein complex of TOM70, Hsp90, IRF3, and the apoptosis regulator Bcl-2-associated X (BAX) protein mediates apoptosis upon viral infection by inducing the burst of cytochrome c into the cytosol (Wei et al., 2015), so ORF9b's inhibitory effect on TOM70-Hsp90 interaction not only disturbs the innate immune pathway but might also modulate apoptosis. Since the induction of apoptosis is an important characteristic of SARS-CoV-2 (Donia & Bokhari, 2021), elucidating the role that ORF9b plays in modulating apoptosis is of great significance and may shed light on the pathogenicity of SARS-CoV-2. Furthermore, since TOM70 has been

demonstrated to be an essential mitochondria import receptor for PTEN-induced kinase 1 (PINK1) responsible for mitochondria quality control (Kato et al., 2013), ORF9b's binding to TOM70 may decrease the transport efficiency of PINK1. Since the accumulation of PINK1 outside the mitochondria is a signal of damaged mitochondria and induces mitophagy (Narendra et al., 2010), ORF9b may cause mitophagy by lowering TOM70 import efficiency of PINK1. Additionally, it has been suggested that localization of SARS-CoV-2 viral proteins to the mitochondria might play crucial roles in SARS-CoV-2 pathogenesis, so understanding the molecular mechanism of ORF9b-TOM70 interaction and its downstream effect on mitochondrial mitophagy would provide important insights into SARS-CoV-2 pathogenesis and offer new perspectives for drug design (Singh et al., 2020). In summary, the effect of ORF9b-TOM70 interaction on inhibiting the immune pathway, modulating apoptosis, and inducing mitophagy should be further investigated to better understand SARS-CoV-2's pathogenicity and provide clues for therapy design. The ORF9b mutations identified in this study that may facilitate its interaction with TOM70 and the postulation of the chaperonic role that Hsp90 plays in ORF9b-TOM70 interaction pave the way for future research on this important interaction essential for understanding SARS-CoV-2 pathogenicity.

8. References

- Bezstarosti K, Lamers MM, Doff WAS, Wever P, Thai K, van Kampen JJ, Haagmans BL, Demmers JA. Targeted proteomics as a tool to detect sars-cov-2 proteins in clinical specimens. 2020. doi:10.1101/2020.04.23.057810
- Biswas S, Shisler JL. Molluscum contagiosum Virus MC159 ABROGATES cIAP1-NEMO interactions and INHIBITS NEMO Polyubiquitination. *Journal of Virology*. 2017;91(15). doi:10.1128/jvi.00276-17
- Bojkova D, Klann K, Koch B, Widera M, Krause D, Ciesek S, Cinatl J, Münch C. Proteomics of sars-cov-2-infected host cells reveals therapy targets. *Nature*. 2020;583(7816):469–472. doi:10.1038/s41586-020-2332-7
- Boni MF, Lemey P, Jiang X, Lam TT-Y, Perry BW, Castoe TA, Rambaut A, Robertson DL. Evolutionary origins of THE SARS-CoV-2 sarbecovirus LINEAGE responsible for the COVID-19 pandemic. *Nature Microbiology*. 2020;5(11):1408–1417. doi:10.1038/s41564-020-0771-4
- Capriotti E, Fariselli P, Casadio R. I-mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research*. 2005;33(Web Server). doi:10.1093/nar/gki375
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An Integrative Toolkit developed for INTERACTIVE analyses of Big biological data. *Molecular Plant*. 2020;13(8):1194–1202. doi:10.1016/j.molp.2020.06.009
- Chen Y, He L, Peng Y, Shi X, Chen J, Zhong J, Chen X, Cheng G, Deng H. The hepatitis C VIRUS Protein Ns3 suppresses Tnf- α -stimulated activation of NF-KB by Targeting I κ B. *Science Signaling*. 2015;8(403). doi:10.1126/scisignal.aab2159
- Crooks GE. WebLogo: A SEQUENCE Logo Generator. *Genome Research*. 2004;14(6):1188–1190. doi:10.1101/gr.849004

- Davidson AD, Williamson MK, Lewis S, Shoemark D, Carroll MW, Heesom KJ, Zambon M, Ellis J, Lewis PA, Hiscox JA, et al. Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Medicine*. 2020;12(1). doi:10.1186/s13073-020-00763-0
- Donia A, Bokhari H. Apoptosis induced by sars-cov-2: Can we target it? *Apoptosis*. 2021;26(1-2):7–8. doi:10.1007/s10495-021-01656-2
- Duhovny D, Nussinov R, Wolfson HJ. Efficient unbound docking of rigid molecules. *Lecture Notes in Computer Science*. 2002:185–200. doi:10.1007/3-540-45784-4_14
- Finkel Y, Mizrahi O, Nachshon A, Weingarten-Gabbay S, Morgenstern D, Yahalom-Ronen Y, Tamir H, Achdout H, Stein D, Israeli O, et al. The coding capacity of sars-cov-2. *Nature*. 2020;589(7840):125–130. doi:10.1038/s41586-020-2739-1
- Fontana J, Fulton D, Chen Y, Fairchild TA, McCabe TJ, Fujita N, Tsuruo T, Sessa WC. Domain mapping studies reveal that the M domain Of hsp90 serves as a molecular scaffold to Regulate Akt-Dependent phosphorylation of Endothelial nitric oxide synthase and no release. *Circulation Research*. 2002;90(8):866–873. doi:10.1161/01.res.0000016837.26733.be
- Gao X, Zhu K, Qin B, Olieric V, Wang M, Cui S. Crystal structure of sars-cov-2 orf9b complex with human tom70. 2021. doi:10.2210/pdb7dhg/pdb
- Gasteiger E. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*. 2003;31(13):3784–3788. doi:10.1093/nar/gkg563
- Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, O'Meara MJ, Rezelj VV, Guo JZ, Swaney DL, et al. A sars-cov-2 protein interaction map reveals targets for drug repurposing. *Nature*. 2020;583(7816):459–468. doi:10.1038/s41586-020-2286-9
- Han L, Zhuang M-W, Zheng Y, Zhang J, Nan M-L, Wang P-H, Gao C. SARS-CoV-2 orf9b Antagonizes Type I and iii Interferons by targeting multiple components of Rig-i/mda-5-mavs, TLR3-TRIF, and Cgas-sting signaling pathways. 2020. doi:10.1101/2020.08.16.252973
- Jiang H-wei, Li Y, Zhang H-nan, Wang W, Yang X, Qi H, Li H, Men D, Zhou J, Tao S-ce. SARS-CoV-2 proteome microarray for global profiling of covid-19 specific IgG and IgM responses. *Nature Communications*. 2020;11(1). doi:10.1038/s41467-020-17488-8
- Jiang H-wei, Zhang H-nan, Meng Q-feng, Xie J, Li Y, Chen H, Zheng Y-xiao, Wang X-ning, Qi H, Zhang J, et al. Sars-cov-2 orf9b suppresses type i interferon responses by targeting tom70. *Cellular & Molecular Immunology*. 2020;17(9):998–1000. doi:10.1038/s41423-020-0514-8
- Jin H-S, Lee D-H, Kim D-H, Chung J-H, Lee S-J, Lee TH. cIAP1, cIAP2, And XIAP Act cooperatively VIA Nonredundant pathways to Regulate Genotoxic Stress–induced NUCLEAR Factor-κB Activation. *Cancer Research*. 2009;69(5):1782–1791. doi:10.1158/0008-5472.can-08-2256
- Jungreis I, Sealfon R, Kellis M. SARS-CoV-2 gene content and COVID-19 Mutation impact by comparing 44 Sarbecovirus genomes. 2020. doi:10.21203/rs.3.rs-80345/v1

- Kato H, Lu Q, Rapaport D, Kozjak-Pavlovic V. Tom70 is essential FOR PINK1 import into Mitochondria. *PLoS ONE*. 2013;8(3). doi:10.1371/journal.pone.0058435
- King A. *Virus taxonomy: Ninth report of the International Committee on taxonomy of viruses*. Oxford: Elsevier; 2012.
- Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research*. 1987;15(20):8125–8148. doi:10.1093/nar/15.20.8125
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. Mega x: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*. 2018;35(6):1547–1549. doi:10.1093/molbev/msy096
- Lackie RE, Maciejewski A, Ostapchenko VG, Marques-Lopes J, Choy W-Y, Duennwald ML, Prado VF, Prado MA. The Hsp70/Hsp90 Chaperone machinery in neurodegenerative diseases. *Frontiers in Neuroscience*. 2017;11. doi:10.3389/fnins.2017.00254
- Lewis BB, Carter RA, Ling L, Leiner I, Taur Y, Kamboj M, Dubberke ER, Xavier J, Pamer EG. Pathogenicity Locus, Core genome, and Accessory Gene contributions To *Clostridium Difficile* virulence. *mBio*. 2017;8(4). doi:10.1128/mbio.00885-17
- Loo Y-M, Gale M. Immune signaling by rig-i-like receptors. *Immunity*. 2011;34(5):680–692. doi:10.1016/j.immuni.2011.05.003
- Lukaszewicz M, Feuermann M, Jérrouville B, Stas A, Boutry M. In vivo evaluation of the context sequence of the translation initiation codon in plants. *Plant Science*. 2000;154(1):89–98. doi:10.1016/s0168-9452(00)00195-3
- Miranda M, Sorkin A. Regulation of receptors and transporters by ubiquitination: New insights into surprisingly similar mechanisms. *Molecular Interventions*. 2007;7(3):157–167. doi:10.1124/mi.7.3.7
- Narendra DP, Jin SM, Tanaka A, Suen D-F, Gautier CA, Shen J, Cookson MR, Youle RJ. PINK1 is Selectively stabilized ON Impaired mitochondria to Activate parkin. *PLoS Biology*. 2010;8(1). doi:10.1371/journal.pbio.1000298
- Offermanns S, Rosenthal W. *Encyclopedia of molecular pharmacology*. Berlin: Springer; 2008.
- Sato S, Fujita N, Tsuruo T. Modulation of akt kinase activity by binding to hsp90. *Proceedings of the National Academy of Sciences*. 2000;97(20):10832–10837. doi:10.1073/pnas.170276797
- Schaecher SR, Pekosz A. SARS coronavirus Accessory gene expression and function. *Molecular Biology of the SARS-Coronavirus*. 2009:153–166. doi:10.1007/978-3-642-03683-5_10
- Schmidt N, Lareau CA, Keshishian H, Ganskih S, Schneider C, Hennig T, Melanson R, Werner S, Wei Y, Zimmer M, et al. The sars-cov-2 rna–protein interactome in infected human cells. *Nature Microbiology*. 2020;6(3):339–353. doi:10.1038/s41564-020-00846-z
- Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Research*. 2005;33(Web Server). doi:10.1093/nar/gki481

- Shi C-S, Qi H-Y, Boularan C, Huang N-N, Abu-Asab M, Shelhamer JH, Kehrl JH. Sars-coronavirus open reading frame-9b suppresses innate immunity by targeting mitochondria and the mavs/traf3/traf6 signalosome. *The Journal of Immunology*. 2014;193(6):3080–3089. doi:10.4049/jimmunol.1303196
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. Fast, scalable generation Of HIGH - QUALITY PROTEIN multiple sequence ALIGNMENTS USING Clustal Omega. *Molecular Systems Biology*. 2011;7(1):539. doi:10.1038/msb.2011.75
- Singh KK, Chaubey G, Chen JY, Suravajhala P. Decoding sars-cov-2 hijacking of host mitochondria IN COVID-19 pathogenesis. *American Journal of Physiology-Cell Physiology*. 2020;319(2). doi:10.1152/ajpcell.00224.2020
- Stacey SN, Jordan D, Williamson AJ, Brown M, Coote JH, Arrand JR. Leaky scanning is the predominant mechanism for translation of human Papillomavirus type 16 E7 Oncoprotein from E6/E7 Bicistronic mRNA. *Journal of Virology*. 2000;74(16):7284–7297. doi:10.1128/jvi.74.16.7284-7297.2000
- Stecher G, Tamura K, Kumar S. Molecular evolutionary genetics analysis (mega) for macos. *Molecular Biology and Evolution*. 2020;37(4):1237–1239. doi:10.1093/molbev/msz312
- Tang ED, Wang C-Y, Xiong Y, Guan K-L. A role for NF- κ B ESSENTIAL MODIFIER/I κ B KINASE- γ (NEMO/IKK γ) Ubiquitination in the activation of the I κ B kinase complex by tumor NECROSIS FACTOR-A. *Journal of Biological Chemistry*. 2003;278(39):37297–37305. doi:10.1074/jbc.m303389200
- The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC. 2011 May 4 [accessed 2021 Jul 30]. <https://bioinformaticschap.wordpress.com/tools/visulization-tools/pymol/>
- V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: Implications For sars-cov-2. *Nature Reviews Microbiology*. 2020;19(3):155–170. doi:10.1038/s41579-020-00468-6
- Wang Y, Cui L, Yang G, Zhan J, Guo L, Chen Y, Fan C, Liu D, Guo D. Hepatitis b e antigen inhibits nf- κ b activity by interrupting k63-linked ubiquitination of nemo. *Journal of Virology*. 2019;93(2). doi:10.1128/jvi.00667-18
- Wei B, Cui Y, Huang Y, Liu H, Li L, Li M, Ruan K-C, Zhou Q, Wang C. Tom70 mediates sendai Virus-Induced Apoptosis on mitochondria. *Journal of Virology*. 2015;89(7):3804–3818. doi:10.1128/jvi.02959-14
- WHO coronavirus (COVID-19) Dashboard. World Health Organization. [accessed 2021 Aug 14]. https://covid19.who.int/?gclid=Cj0KCQjwI-SHBhCQARIsAFIFRVU2r-8z9ushd18PTIYQ5jUOK4lsWprFjWQOL_ml1BiI-GHnej5rwlgaAl3pEALw_wcB
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview version 2--A multiple sequence Alignment editor and ANALYSIS WORKBENCH. *Bioinformatics*. 2009;25(9):1189–1191doi:10.1093/bioinformatics/btp033
- Wu F. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579:265–269. doi:10.1038/s41586-020-2008-3

- Wu J, Shi Y, Pan X, Wu S, Hou R, Zhang Y, Zhong T, Tang H, Du W, Wang L, et al. Sars-cov-2 orf9b inhibits rig-i-mavs antiviral signaling by interrupting k63-linked ubiquitination of nemo. *Cell Reports*. 2021;34(7):108761. doi:10.1016/j.celrep.2021.108761
- Wu J, Shi Y, Pan X, Wu S, Hou R, Zhang Y, Zhong T, Tang H, Du W, Wang L, et al. Sars-cov-2 orf9b inhibits rig-i-mavs antiviral signaling by interrupting k63-linked ubiquitination of nemo. *Cell Reports*. 2021;34(7):108761. doi:10.1016/j.celrep.2021.108761
- Xu K, Zheng B-J, Zeng R, Lu W, Lin Y-P, Xue L, Li L, Yang L-L, Xu C, Dai J, et al. Severe acute respiratory syndrome coronavirus accessory protein 9b is a virion-associated protein. *Virology*. 2009;388(2):279–285. doi:10.1016/j.virol.2009.03.032
- Xu K, Zheng B-J, Zeng R, Lu W, Lin Y-P, Xue L, Li L, Yang L-L, Xu C, Dai J, et al. Severe acute respiratory syndrome coronavirus accessory protein 9b is a virion-associated protein. *Virology*. 2009;388(2):279–285. doi:10.1016/j.virol.2009.03.032
- Young JC, Moarefi I, Hartl FU. Hsp90. *Journal of Cell Biology*. 2001;154(2):267–274. doi:10.1083/jcb.200104079
- Young JC, Obermann WMJ, Hartl FU. Specific binding of tetratricopeptide repeat proteins to The C-terminal 12-kda domain Of hsp90. *Journal of Biological Chemistry*. 1998;273(29):18007–18010. doi:10.1074/jbc.273.29.18007

9. Supplementary Material

The mutation frequency of the ORF9b amino acid position 43-78 that interacts with the C-terminal domain of TOM70 was calculated in the software Jalview and recorded as a percentage in the table below. Table 1-7 contains the timeline of mutation frequency of each amino acid position from January 2020 to June 2021

Table 1. Mutation frequency timeline of ORF9b amino acid position 43-47

43		44		45		46		47	
20May	P43L:0.03	20Mar	I44T:0.01	20July	M45L:0.03	20July	L46R:0.01	20July	R47L:0.01
20July	P43L:0.01	20Dec	I44H:0.02	20Aug	M45L:0.22	20Aug	L46R:0.02	20Aug	R47L:0.02
20Aug	P43L:0.02	21Jan	I44N:0.01	20Sep	M45L:0.39	20Dec	L46A:0.02	20Oct	R47L:0.02
20Dec	P43L:0.04		M44*:0.01	20Nov	M45V:0.03	21Jan	L46A:0.01	20Dec	R47S:0.02
	P43H:0.05	21Mar	I44N:0.005	20Dec	M45T:0.04		L46C:0.01		R47K:0.02
21Jan	P43L:0.03		M4H:0.005		M45V:0.02	21Feb	L46P:0.006	21Jan	R47H:0.04
	P43N:0.01	21Apr	M4D:0.009	21Mar	M45T:0.01	21Mar	L46A:0.005		R47S:0.01
	P43Q:0.01		M4L:0.006		M45L:0.005		L46C:0.005		R47V:0.01
21Feb	P43L:0.004		M4Y:0.003		M45V:0.005	21Apr	L46C:0.003	21Feb	R47H:0.006
	P43S:0.006	21May	M4*:0.002	21Apr	M45V:0.009		L46C:0.003	21Mar	R47C:0.015
21Mar	P43L:0.02	21June	M4*:0.04		M4L:0.005		L46P:0.003		R47L:0.01
	P43N:0.005				M45T:0.005	21May	L46C:0.009		R47S:0.01
	P43S:0.005				M4Y:0.004		L46P:0.004	21Apr	R47Q:0.003
21Apr	P43L:0.012			21May	M45V:0.009		M45V:0.004		R47H:0.003
	P43Q:0.003				M4L:0.004	21June	L46C:0.04		M47H:0.003
	P43S:0.003			21June	M45Y:0.04			21May	R47H:0.009
21May	P43L:0.018				M45T:0.02				R47V:0.009
	P43Q:0.009								R47L:0.004
21June	P43Q:0.04								R47K:0.004
	P43L:0.02							21June	R47C:0.04
									R47L:0.04
									R47V:0.04

Table 2. Mutation frequency timeline of ORF9b amino acid position 48-52

48		49		50		51		52	
20July	L48G:0.01	20July	G49S:0.01	20Mar	S50L:0.01	20June	P51L:0.03	20Mar	L52K:0.01
20Aug	L48G:0.02	20Aug	G49S:0.02	20Apr	S50L:0.02	20July	P51L:0.04	20July	L52S:0.01
	L48P:0.02	20Dec	G49P:0.02	20May	S50L:0.03	20Aug	P51L:0.09	20Aug	L52P:0.02
20Dec	L48W:0.02	21Jan	G49P:0.01	20June	S50T:0.03	20Dec	P51A:0.02		L52S:0.02
21Jan	L48H:0.01		G49Y:0.01		S50C:0.03		P51L:0.02	20Oct	L52P:0.05
21Mar	L48W:0.005	21Mar	G49P:0.005	20July	S50L:0.01	21Jan	P51L:0.03	21Jan	L52S:0.01
21Apr	L48G:0.003		G49I:0.005		S50P:0.01		P51A:0.01	21Mar	L52N:0.005
	L48H:0.003	21Apr	G49L:0.003	20Aug	S50L:0.02		P51R:0.01		L52Q:0.005
	L48P:0.003		G49H:0.003		S50P:0.02		P51K:0.01	21Apr	L52S:0.009
21May	L48G:0.004		G49P:0.003	20Sep	S50L:0.05	21Feb	P51L:0.065		L52V:0.003
		21May	G49V:0.009	20Oct	S50L:0.05		P51K:0.016	21May	L52S:0.013
			G49S:0.024	20Nov	S50L:0.03	21Mar	P51L:0.124	21June	L52S:0.04
		21June	G49V:0.04	20Dec	S50L:0.04		P51Q:0.01		
					S50P:0.02		P51A:0.005		
				21Jan	S50L:0.15	21Apr	P51L:0.113		
					S50T:0.02		P51R:0.003		
					S50P:0.01		P51K:0.003		
				21Feb	S50L:0.097	21May	P51H:0.009		
					S50T:0.005		P51L:0.009		
				21Apr	S50L:0.063	21June	P51L:0.017		
					S52H:0.003		P51R:0.04		
					S50P:0.003				
				21May	S50L:0.057				
					S50R:0.009				
					S50P:0.004				
				21June	S50L:0.11				
					S50K:0.04				

Table 3. Mutation frequency timeline of ORF9b amino acid position 53-57

	S3		S4		S6		S8		S7
20July	S53L:0.01	20July	L54H:0.01	20Apr	N55K:0.02	20Mar	M56I:0.01	20Mar	A57V:0.02
20Aug	S53L:0.02	20Aug	L54N:0.02	20July	N55M:0.01	20Apr	M56T:0.02	20Apr	A57V:0.02
20Dec	S53L:0.02	20Dec	L54Q:0.02		N55K:0.01	20May	M56I:0.03	20May	A57V:0.034
	S53T:0.02	21Jan	L54Q:0.01	20Aug	N55M:0.02	20June	M56I:0.06	20July	A57V:0.03
21Jan	S53H:0.01		L54R:0.01	20Nov	N55S:0.03	20July	M56A:0.01		A57R:0.01
	S53T:0.01		L54S:0.01	20Dec	N55H:0.02		M56I:0.01	20Aug	A57V:0.02
21Feb	S53L:0.032	21Feb	L54K:0.009	21Jan	N55H:0.01	20Aug	M56A:0.02		A57R:0.02
21Mar	S53M:0.005	21Mar	L54A:0.005		N55T:0.01	20Sep	M56I:0.1	20Oct	A57V:0.05
	S53T:0.005		L54Q:0.005	21Feb	N55S:0.008	20Dec	M56G:0.02	20Nov	A57V:0.17
21Apr	S53L:0.024	21Apr	L54A:0.003	21Mar	N55S:0.01	21Jan	M56G:0.01		A57E:0.03
	S53H:0.003		L54N:0.003		N55R:0.005		M56W:0.01	20Dec	A57G:0.5
	S53M:0.003		L54S:0.003		N55R:0.005	21Mar	M56G:0.005		A57V:0.14
21May	S53H:0.009	21May	L54S:0.009	21Apr	N55S:0.006		M56K:0.005		A57K:0.14
	S53L:0.009		L54N:0.004		N55M:0.003	21Apr	M56I:0.006		A57K:0.02
21June	S53H:0.04		L54P:0.004		N55R:0.003		M56A:0.003	21Jan	A57V:0.19
	S53L:0.02	21June	L54S:0.04		N55T:0.003		M56K:0.003		A57K:0.01
				21May	N55T:0.009		M56T:0.002		A57Q:0.01
					N55M:0.004		M56W:0.003	21Feb	A57V:0.202
				21June	N55T:0.03	21May	M56W:0.009		A57K:0.016
							M56A:0.004	21Mar	A57V:0.164
						21June	M56W:0.04		A57K:0.2
									A57K:0.005
									A57T:0.005
								21Apr	A57V:0.164
									A57Q:0.003
									A57T:0.003
									A57S:0.003
								21May	A57V:0.089
									A57Q:0.009
									A57R:0.004
								21June	A57V:0.05
									A57Q:0.04

Table 4. Mutation frequency timeline of ORF9b amino acid position 58-62

	S8		S9		S9		S11		S12
20May	R58S:0.09		K59N:0.01	20May	R60A:0.08	20July	L61N:0.01	20June	R62S:0.03
20June	R58K:0.03		K59P:0.0	20July	R60L:0.01	20Aug	L61N:0.02	20July	R62S:0.011
20July	R59Q:0.03	20May	K59P:0.171	20Aug	R60L:0.02	20Nov	L61S:0.03	20Aug	R62S:0.02
	R58K:0.01	20July	K59N:0.01	20Dec	R60L:0.02	20Dec	L61K:0.02	20Sep	R62T:0.06
20Aug	R58Q:0.02		K59T:0.01	21Jan	R60L:0.01	21Jan	L61K:0.01	20Dec	R62P:0.02
20Dec	R58K:0.02	20Aug	K59T:0.02		R60P:0.01	21Feb	L61K:0.008	21Jan	R62P:0.01
	R58E:0.02	20Dec	K59R:0.04		R60Q:0.01	21Mar	L61K:0.005		R62Q:0.01
21Jan	R58E:0.01		K59D:0.02	21Feb	R60K:0.006	21Apr	L61N:0.003	21Feb	R62Q:0.016
21Feb	R58K:0.016		K59K:0.02	21Mar	R60A:0.01		L61L:0.002		R62S:0.008
21Mar	R58K:0.1	21Jan	K59P:0.05		R60Q:0.01	21May	L61L:0.008	21Mar	R62D:0.025
	R58E:0.005		K59Q:0.01		R60L:0.003		L61N:0.004		R62S:0.015
	R58L:0.005		K59N:0.01		R62S:0.009	21June	L61L:0.04		R62E:0.005
21Apr	R60Q:0.007		K59Q:0.01	21Apr	R60A:0.043				R62I:0.005
	R58N:0.005	21Feb	K59N:0.04		R60L:0.015			21Apr	R62D:0.001
	R60M:0.006		K59K:0.016		R60L:0.003				R62E:0.003
	R58L:0.003	21May	K59R:0.035		R60P:0.005				R62H:0.009
	R58T:0.003		K59Q:0.005		R62S:0.003				R62H:0.003
21May	R58K:0.015	21Apr	K59R:0.024		R60V:0.003				R62I:0.003
	R59Q:0.001		K59R:0.006	21May	R60A:0.002				R62Q:0.003
21June	R58Q:0.04		K59M:0.003		R60P:0.009			21May	R62D:0.004
	R58K:0.02		K59T:0.002		R60Q:0.004				R62I:0.001
		21May	K59N:0.025		R60L:0.004				R62S:0.004
			K59R:0.018	21June	R60A:0.07			21June	R62D:0.13
			K59K:0.009		R60E:0.07				R62Q:0.04
			K59T:0.004		R60P:0.04				
		21June	K59N:0.04		R60K:0.02				
			K59H:0.04						

Table 5. Mutation frequency timeline of ORF9b amino acid position 63-67

63		64		65		66		67	
20June	S63K:0.03	20Mar	L64X:0.01	20July	E65D:0.01	20Mar	D66V:0.04	20June	K67T:0.03
20July	S63L:0.01	20Apr	L64X:0.02	20Aug	E65D:0.02	20May	D66V:0.029	20July	K67A:0.01
20Aug	S63L:0.03	20May	L64X:0.009	20Oct	E ⁺ :0.05	20July	D66G:0.01	20Aug	K67A:0.02
20Sep	S63X:0.05	20July	L64X:0.01	20Nov	E65A:0.03		D66K:0.01	20Sep	K67R:0.05
20Nov	S63X:0.06		L64E:0.01	20Dec	E65G:0.02	20Aug	D66G:0.02	20Oct	K67E:0.05
20Dec	S63X:0.21	20Aug	L64E:0.02	21Jan	E65G:0.01		D66K:0.02	20Nov	K67E:0.06
	S63P:0.02	20Dec	L64V:0.09		E65K:0.01	20Dec	D66Q:0.02	20Dec	K67G:0.02
21Jan	S63X:0.13		L64F:0.04		E65R:0.01	21Jan	D66Q:0.01	21Jan	K67G:0.01
	S63L:0.01		L64R:0.04	21Feb	E65Q:0.008		D66T:0.01		K67R:0.01
21Feb	S63X:0.121	21Jan	L64F:0.03	21Mar	E65A:0.005	21Mar	D66F:0.005	21Feb	K67E:0.032
	S63F:0.008		L64V:0.03		E65G:0.005		D66Q:0.005	21Mar	K67G:0.005
21Mar	S63F:0.025		L64R:0.01		E65K:0.005	21Apr	D66N:0.009		K67Q:0.005
	S63X:0.015		L64S:0.01		E65X:0.005		D66F:0.003	21Apr	K67A:0.003
	S63D:0.005	21Feb	L64F:0.097	21Apr	E65A:0.003		D66K:0.003		K67Q:0.003
	S63P:0.005		L64V:0.032		E65D:0.003		D66T:0.003		K67R:0.003
21Apr	S63F:0.024		L64I:0.016		E65R:0.003	21May	D66T:0.009	21May	K67R:0.022
	S63D:0.003	21Mar	L64F:0.07	21May	E65R:0.009		D66G:0.004		K67A:0.004
	S63L:0.003		L64I:0.005		E65D:0.004		D66K:0.004	21June	K67R:0.07
	S63P:0.003		L64K:0.005		E65V:0.004	21June	D66T:0.04		
	S63K:0.003		L64R:0.005	21June	E65R:0.04		D66A:0.02		
21May	S63F:0.009		L64X:0.005		E65G:0.02				
	S63P:0.009	21Apr	L64F:0.14		E65K:0.02				
	S63L:0.004		L64P:0.006						
	S63X:0.004		L64E:0.003						
21June	S63P:0.04		L64K:0.003						
	S63F:0.02		L64S:0.003						
		21May	L64F:0.125						
			L64P:0.009						
			L64S:0.009						
			L64E:0.004						
			L64X:0.004						
		21June	L64F:0.24						
			L64S:0.04						

Table 6. Mutation frequency timeline of ORF9b amino acid position 68-72

68		69		70		71		72	
20May	A68V:0.017	20July	F69C:0.01	20July	Q70L:0.01	20July	L71T:0.01	20May	T72I:0.09
20July	A68F:0.01	20Aug	F69C:0.02	20Aug	Q70L:0.02	20Aug	L71T:0.02	20June	T72I:0.06
	A68S:0.01	20Dec	F69P:0.02	20Oct	Q70L:0.05	20Dec	L71N:0.02	20July	T72P:0.01
20Aug	A68F:0.02	21Jan	F69H:0.01	20Nov	Q70L:0.03	21Jan	L71N:0.01	20Aug	T72I:0.02
	A68V:0.02		F69S:0.01	20Dec	Q70I:0.02		L71 ⁺ :0.01		T72S:0.02
	A68X:0.03	21Mar	F69P:0.005		Q70L:0.02	21Feb	L71K:0.008	20Nov	T72I:0.06
20Dec	A68V:0.02		F69T:0.005	21Jan	Q70I:0.01			20Dec	T72I:0.02
21Jan	A68V:0.04	21Apr	F69C:0.003		Q70N:0.01			21Jan	T72I:0.07
	A68R:0.01		F69S:0.003	21Mar	Q70L:0.01	21Mar	L71I:0.005	21Feb	T72I:0.024
21Feb	A68V:0.016		F69T:0.003		Q70I:0.005		L71N:0.005	21Mar	T72A:0.005
21Mar	A68V:0.015		F69Y:0.003		Q70P:0.005	21Apr	L71I:0.003		T72I:0.005
	A68V:0.005	21May	F69S:0.009	21Apr	Q70L:0.024		L71I:0.003	21Apr	T72I:0.012
21Apr	A68V:0.012		F69C:0.004		Q70H:0.006		L71 ⁺ :0.003		T72A:0.003
	A68F:0.003		F69Y:0.004		Q70N:0.003	21May	L71 ⁺ :0.009		T72H:0.003
	A68L:0.003	21June	F69S:0.04		Q70P:0.003		L71T:0.004		T72P:0.003
	A68R:0.003			21May	Q70L:0.027	21June	L71 ⁺ :0.04	21May	T72H:0.009
21May	A68V:0.031				Q70R:0.022				T72I:0.004
	A68R:0.009				Q70N:0.009				T72P:0.004
	A68F:0.004			21June	Q70N:0.01			21June	T72H:0.04
21June	A68R:0.04				Q70L:0.02				T72I:0.02

Table 7. Mutation frequency timeline of ORF9b amino acid position 73-78

73		74		75		76		77		78	
20Mar	P73S:0.01	20June	I74L:0.03	20Mar	A75K:0.01	20June	V76P:0.03	20Apr	Q77H:0.04	20Mar	M78T:0.01
20May	P73S:0.009	20July	I74A:0.01	20May	A75V:0.009	20July	V76Q:0.01		Q77:-0.02	20May	M78T:0.009
20June	P73S:0.03	20Aug	I74A:0.02		A75X:0.009	20Aug	V76Q:0.02	20May	Q77H:0.017	20June	M78A:0.03
	P73Q:0.03	20Dec	I74S:0.02	20June	A75T:0.03	20Nov	V76L:0.03	20June	Q76E:0.03	20July	M78T:0.03
20July	P73I:0.01		I74T:0.02	20July	A75V:0.03	20Dec	V76P:0.02	20July	Q77H:0.13	20Aug	M78I:0.09
20Aug	P73I:0.02	21Jan	I74M:0.01	20Aug	A75V:0.02	21Jan	V76P:0.01		Q77M:0.01		M78T:0.02
	P73S:0.02		I74S:0.01	20Oct	A75V:0.05		V76S:0.01	20Aug	Q77M:0.02	20Nov	M78T:0.06
20Sep	P73L:0.05		I74:-0.01	20Dec	A75K:0.04	21Feb	V76L:0.016	20Sep	Q77X:0.1	20Dec	M78D:0.02
20Oct	P73S:0.05	21Mar	I74M:0.005		A75S:0.02	21Mar	V76L:0.02		Q77E:0.05		M78T:0.02
20Dec	P73N:0.02		I74Q:0.005		A75V:0.02		V76P:0.005		Q77H:0.05	21Jan	M78T:0.07
21Jan	P73N:0.01		I74T:0.005	21Jan	A75V:0.03		V76T:0.005	20Oct	Q77:-0.05		M78D:0.01
21Mar	P73N:0.005		I74S:0.005		A75V:0.01	21Apr	V76L:0.003	20Nov	Q77E:0.03		M78V:0.01
	P73V:0.005	21Apr	I74A:0.003		A75S:0.01		V76Q:0.003	20Dec	Q77X:0.04		M78:-0.01
21Apr	P73I:0.003		I74Q:0.003	21Feb	A75V:0.032		V76S:0.003		Q77D:0.02	21Feb	M78V:0.008
	P73L:0.003		I74:-0.003	21Mar	A75V:0.03		V76T:0.003		Q77E:0.02		M78:-0.008
	P73Q:0.003	21May	I74M:0.018		A75M:0.005	21May	V76F:0.013		Q77H:0.02	21Mar	M78T:0.015
	P73V:0.003		I74:-0.009		A75S:0.005		V76S:0.009		Q77:-0.02		M78D:0.005
21May	P73S:0.049		I74A:0.004	21Apr	A75V:0.095		V76Q:0.004	21Jan	Q77E:0.09		M78L:0.005
	P73L:0.009	21June	I74:-0.04		A75M:0.003	21June	V76S:0.04		Q77H:0.03		M78L:0.005
	P73Q:0.009				A75Q:0.003				Q77K:0.02	21Apr	M78T:0.018
	P73I:0.004			21May	A75V:0.089				Q77D:0.01		M78L:0.003
21June	P73Q:0.04				A75Q:0.009				Q77R:0.01		M78:-0.003
					A75X:0.004			21Feb	Q77E:0.145	21May	M78:-0.009
				21June	A75V:0.09				Q77:-0.024		M78T:0.004
					A75Q:0.04				Q77H:0.016	21June	M78:-0.04
								21Mar	Q77E:1.306		M78T:0.02
									Q77K:0.02		
									Q77H:0.015		
									Q77:-0.01		
									Q77D:0.005		
								21Apr	Q77E:3.909		
									Q77H:0.045		
									Q77K:0.006		
									Q77:-0.006		
									Q77G:0.003		
									Q77G:0.003		
									Q77L:0.003		
									Q77M:0.003		
									Q77P:0.003		
									Q77R:0.003		
									Q77X:0.03		
								21May	Q77E:5.899		
									Q77H:0.049		
									Q77K:0.013		
									Q77R:0.013		
									Q77G:0.004		
									Q77M:0.004		
									Q77:-0.004		
								21June	Q77E:13.34		
									Q77R:0.04		
									Q77A:0.02		
									Q77X:0.02		



Expanding Intergovernmental Water Markets Between Colorado River Basin Tribal Countries and Western States as a Response to Socioeconomic Drought

Xinran (Cici) Xu

Author Background: *Xinran (Cici) Xu grew up in China and the United States and currently attends Saratoga High School in Saratoga, California in the United States. Her Pioneer research concentration was in the field of environmental studies and titled “Environmental Justice and Resource Conflicts.”*

Abstract

Drought is an ongoing issue in western North America, exacerbated in frequency and severity by global climate change. In recent years, the Colorado River has been experiencing a critical socioeconomic drought where water demand exceeds supply. This puts the basin’s current water management policies on the federal level under the spotlight as many policies inadequately address the severity of drought. Conflicts over water resources arose as water appropriators like Arizona farmers were impacted more acutely by the decrease of water allocations from the Colorado River while other water appropriators like tribal countries around the river experienced annual underutilization of their water rights due to financial constraints. The perspectives of Arizona farmers and tribal nations were often marginalized and not thoroughly represented in discussions of droughts, which demonstrates a lack of representation and ability to negotiate or make decisions about their access, use, and control over water. This research paper analyzes the existing water governmental system of the Colorado River as well as individual tribal nations, identifies flaws within each system by referencing Elinor Ostrom’s *Governing the Commons*, applies past case studies that successfully addressed water shortages and establishment of water markets, and proposes a concrete structure of the Colorado River Basin (CRB) intergovernmental water market between the tribal nations and the states within the United States.

1. Introduction

California is experiencing the worst drought in 1,200 years, caused by global climate change and growing demand for water in the west (Borunda, 2021). The water level of Lake Mead, the Lower Basin of the Colorado River System and the largest reservoir in the United States dropped to 143 feet below the water level in 1930 when

it was classified as “full,” according to the US Bureau of Reclamation, the primary manager of surface water resources (Edalat and Stephen 2019). Crucial to the storage and supply of clean drinking water to the southwest region of the United States in dry seasons, Lake Mead’s continuous diminishment of freshwater poses risks to around 40 million people in Arizona, California, Wyoming, Colorado, New Mexico, Utah, Nevada, and Mexico (see Figure 1) (Edalat and Stephen 2019). As Lake Mead’s socioeconomic drought continues where water demand exceeds water supply, it prompts the water basin system to become increasingly vulnerable. This increase in unforeseen vulnerability poses a risk of future conflicts between appropriators, such as farmers, tribal communities, and other municipal and industrial water users. Critical reflections upon the existing CRB management policies and its historical approach to water management are needed to assess their limitations and drive new institutional change. This action would better achieve equity in water usage and economic development in the CRB region and a higher level of collective action between water rights holders.

The CRB is often known as one of the most over-allocated water systems in the world with strict volume control and a clear division of water allocation rights (Edalat and Stephen 2019). The CRB system is managed by a Leviathan approach in which an external central authority obtains absolute control over the appropriation of water rights among the seven states as well as sovereign tribal countries. When the Bureau of Reclamation was established in 1902, the federal government was pressured to undertake irrigation projects in the western states as a response to the lack of precipitation and the growing demand for water due to droughts (the Bureau of Reclamation). Because the former state, local and private level irrigation projects had failed to effectively conserve water using adequate technologies, this need-based action of the federal government laid the basis for future water management policies that pursue the goal of “reclamation,” which is historically defined as irrigation projects to reclaim land for human use (the Bureau of Reclamation).

1.1. Background of State CRB Water Allocation

Following the establishment of the Reclamation came a century of compacts, agreements, and treaties, known as the *Law of the River*, created to address and resolve conflicts over the intercity, interstate, and international water allocations, which demonstrated a gradual trend of increasing federal agencies’ localized interference in detailing the specific water rights obtained by each state and water district (CRS, 2021). In 1922, based on the average historic annual flows of 16.4 MAF, the Colorado River Compact (CRC) assigned the appropriation of water between the Lower and Upper Basin of the CRB for the first time, giving each basin 7.5 million acre-feet (MAF) of water while granting the Lower Basin the leeway to exceed 1 MAF of usage (CRS, 2021). In 1928, the Boulder Canyon Project Act (BCPA) modified the CRC of 1922, initiating a federal project on water facilities such as the Hoover Dam to transfer water to Southern California (CRS, 2021). The BCPA created Lake Mead with a capacity of 26.1 MAF water and divided the 7.5 MAF Lower Basin appropriation more specifically among the three states, with California holding 4.4 MAF, Arizona 2.8 MAF, Nevada 0.3 MAF, and the surplus split equally among the three states. A continuation of BCPA, the California Seven Party Agreement of 1931 clarified and assigned specific amounts of water to the seven water districts in California to settle water rights conflicts over the allocation of the Colorado River System (Figure 1) (CRS, 2021). The detailed allocation efforts of the Upper CRB came later in 1948, assigning a specific percentage of the 7.5 MAF of

water to Upper Basin states (Colorado: 51.75%; New Mexico: 11.25 %; Utah: 23%; Wyoming: 14%; Upper Arizona: 50,000 acre-feet (CRS, 2021).

However, conflicts arose between Arizona and California in 1963 as CRB water rights were directed by the system of *prior appropriation* that assigns seniority of water rights based on who puts the water to beneficial use (irrigation or industrial projects like mining) first and not based on land ownership (CRS, 2021). This “first-come-first-serve” institutional rule dictated the conflicts among states and water rights holders in the CRB. For example, in the mid-19th century, California was afforded greater advantages over water allocations because of its economic development and population growth. When Arizona wanted to establish projects that transfer CRB tributary water to Tucson and Phoenix, California used its seniority in *prior appropriation* to oppose it. This conflict was settled by the Supreme Court case *Arizona v. California*, replacing the prior appropriation system with BCPA water rights and authorizing future projects such as the Central Arizona Project (CAP) within the Colorado River Basin Project Act to deliver water to southern Arizona using the 336-mile delivery system with the condition of prioritizing California’s 4.4 MAF designated water rights in times of drought (CRS, 2021). Therefore, the CAP holds junior entitlements to water allocation. This court case also delegates the Secretary of Interior as the “water master,” overlooking the states’ rights over CRB water allocations.

As the water consumption increased dramatically between 1971 and 2002, the annual river flow decreased to 12.4 MAF (CRS, 2021). New CRB management policies have been authorized by Congress in response to the increasingly severe drought. The jurisdiction of CAP began to manifest in the 2007 Interim Guidelines/Coordinated Operations for Lake Powell and Lake Mead, otherwise known as the 2007 Drought Contingency Plan (DCP), which authorized conservation of water in Lake Mead to create Intentionally Created Surplus (ICS). The 2007 DCP reduced allocation from Arizona’s CAP and Nevada, as when the water level of Lake Mead dropped below 1075 ft. This change in allocation created an immediate impact on the Arizona farmers’ community, cutting 0.3 AF of irrigation water from Arizona agriculture, which led more farmers to base their water use on self-abstracted groundwater (CRS, 2021).

1.2. Background of Sovereign Tribal Countries’ CRB Water Allocation

Tribal water rights in the Colorado River Basin have a long history of dispute by the public and the federal government. In 1908, the Winters Doctrine, a result of the *Winters v. the United States* Supreme Court case, initiated the process of settling disputes over tribal water rights (CRS, 2021). Under the doctrine, the federal government of the United States implied the acknowledgment of water rights of tribal reservations so they could be self-reliant and self-sufficient as an effort to mitigate disputes over the diversion of water from Milk River, a major source of irrigation water for the Fort Belknap Indian Reservation in Montana (the Bureau of Reclamation, 2018). It was not until 1963 in the *Arizona v. California* Supreme Court case that the federal government explicitly established that Native American tribes hold senior rights in the prior appropriation system. This jurisdiction establishes the “practicably irrigable acreage” (PIA) that quantifies tribal communities’ water claims for agricultural purposes (the Bureau of Reclamation, 2018). Moreover, an alternative quantification approach, “homeland standard,” assesses tribal communities’ general water needs (including evidence from the past, present, and future) to determine the amount of water allocation needed for each tribe (the Bureau of Reclamation, 2018).

Both the *Winters v. United States* and the *Arizona v. California* Supreme Court cases demonstrated progress in considering the role of tribal communities in water management policy while leaving tribal perspectives passive in the discussion. While these legal decisions certainly acknowledge and allocate water rights to tribal Nations, they resulted in common misconceptions about the legal status of tribal communities among the public. Tribal communities are sovereign nations while adhering to federal law. Therefore, though tribal communities are scattered in states and the joint borders between states, they are not subject to state regulation.

However, the CRB tribal communities are becoming a stronger force to be reckoned with as tribal water settlement claims are being continuously ratified. Specific to Arizona and its relationship with CRB tribal communities, 46% of the CAP water supply goes to the 22 Arizona Indian Tribes (CAP, 2021). The 2004 Arizona Water Settlements Act acknowledged Arizona, the Gila River Indian Community, and the Tohono O'odham Nation as water rights settlements (the Bureau of Reclamation, 2004). Moreover, it advanced tribal water rights in the form of new tribal and non-tribal allocation of CAP water to achieve 50% of the allocation to Native American communities and allowed tribal communities to lease water within the state and under the Arizona Water Banking Authority (CAP, 2021).

In times of water supply shortage, tribal communities obtain an increasingly critical role in mitigating the effects of drought in tribal as well as non-tribal communities, as they collectively hold 2.8 MAF of CRB water (the Bureau of Reclamation, 2018). Yet, many tribes are unable to fully recognize and utilize their water rights (the Bureau of Reclamation, 2018). The 2018 Tribal Water Study, conducted jointly by the Bureau of Reclamation and the Colorado River Basin Ten Tribes Partnership, expresses a common challenge faced by Partnership Tribes that hinders them from acquiring better water delivery facilities and water usage and supply data collection monitors: lack of permanent funding. According to the study, the lack of funding stems from the lack of public recognition about the importance of tribal water rights in the CRB, therefore resulting in a lack of support (the Bureau of Reclamation, 2018). Though some tribal nations like Quechan Indian Tribe and Jicarilla Apache Nation can get approvals from the Secretary of Interior and the government to lease water to other non-tribal entities and earn revenues, these opportunities are not currently available to all tribal communities (the Bureau of Reclamation, 2018).

In this study, I aim to evaluate the limitations of CRB centralized management and how this form of governance hinders CRB communities in approaching equity in water allocations during drought. To mitigate this issue and strengthen collective resiliency that allows a region to adjust, survive, and adapt to environmental disasters like drought, expanding intergovernmental water markets between CRB tribal communities like Ten Tribes Partnership with the state of Arizona will generate consistent revenues for tribal countries to stimulate their economic developments while satisfying the water demand of the Arizona farmers in agricultural productions. To achieve this goal, two case studies are developed, one in Northern Greece of Chalkidiki and the other in Australia, to demonstrate the potential of water markets in responding to regional water shortages. By connecting the successes and failures from both case studies, an intergovernmental water market is proposed. Ultimately, this study concludes that expanding water markets to embrace more tribal communities is a beneficial and accessible solution to the CRB socioeconomic drought, despite fundamental changes to the existing governance structure of the river.



Figure 1. The map of the Colorado River from the USGS public domain, indicating Lees Ferry as the dividing point between the Upper and Lower Basin and the scale of CRB as being a highly complex interstate and intercountry system. (Source: Public Domain)

2. Literature Review

By acknowledging the complexity of self-governing common-pool resources (CPRs) and the limitations surrounding simple associations between an institutional problem with a unifying solution like “privatization always resolves unregulated uptakes of public resources,” Elinor Ostrom in her work *Governing the Commons: The Evolution of Institutions for Collective Action* critiques the three influential models of governing—the tragedy of the commons, the prisoner’s dilemma game, and the logic of collective action—fundamental to the analysis of CPR like the CRB (Ostrom, 2019). Though the tragedy of the commons was brought to light by Garret Hardin in his *Science* article in 1968, this idea has been continuously reflected upon throughout history by various philosophers and economists as the basis of politics: humans have insatiable desires to selfishly obtain public goods, which result in societal conflicts (Ostrom, 2019). This model has been used widely in the analysis of overpopulation, grazing, and fishery metaphors. The prisoner’s dilemma game provides a related perspective in which the “prisoners” are given the option of cooperating with each other to minimize both sides’ “jail time” or deflecting so that one side takes the complete blame and serves extended jail time while the other goes free (Ostrom, 2019). This model stimulates “prisoners” to always deflect for the potential immediate benefits for themselves. This mentality can be reflected through similar empirical situations in which if both users of the CPRs cooperate, the benefits for both can be maximized; if the users of the CPR do not cooperate and both decide to deflect, no benefits will be granted to either side (Ostrom, 2019). The third model is the logic of collective action in which all sides voluntarily work together to achieve

the maximum benefits (Ostrom, 2019). Though the three models dominate the categorization of policies and can be reflected in empirical policies governing the CRB, it is crucial to recognize their limitations and think outside the box (Ostrom, 2019). Metaphorical analysis of policy frequently and unconsciously strengthens the mentality that the empirically impacted communities and people are “helpless prisoners” who have no power to change the rules. However, this mentality is not true in real life, as a centralized system like CRB management can evolve to better integrate previously isolated but now increasingly powerful perspectives such as those of tribal communities. Without realizing the limitations imposed by these models, generalized governance like Leviathan or privatization can hinder the expansion of innovative ideas such as internal and voluntary self-enforcement of rules (Game 5) (Ostrom, 2019). To further elaborate on these theoretical rules, Ostrom delves into case studies of failure and success in governing CPRs. One case study that demonstrates the institutional change is the competitive pumping races in Southern California that detail the litigations in the three groundwater basins—the Raymond Basin, the West Basin, and the Central Basin. The conflicts were later resolved through voluntary associations of joint discussions, which can be connected to the CRB management dilemma (Ostrom, 2019).

Effective voluntary and joint actions between appropriators in the Colorado River region are critical at this point. Many previous studies have been dedicated to assessing the effect of climate change on Colorado River Basin streamflow. According to *the State of America’s Water Resource*, climate change disrupts traditional precipitation patterns; raises sea levels as it warms and melts polar sea and land ice, causing saltwater intrusion into freshwater; and exacerbates runoff of contaminants into freshwater, degrading its quality (McNabb, 2017). The impacts of climate change are worsened by the population migration from the north and northeastern states to the South and West. Barnett and Pierce (2008) conclude that if global climate change prompts a 20% reduction of water storage due to runoff, Lake Mead is estimated to go dry in 2028. However, this research also indicates that curtailing consumption by 25% can make a dramatic difference in maintaining sustainable water storage. This poses stark challenges for the future as water management policy has to be continuously updated to adapt to the reduced supply of water, which will exacerbate environmental injustice among communities, specifically Arizona farmers and indigenous communities, as revealed through the past case study on the degradation of Klamath River and its devastating impacts on the Karuk Tribe (Fitchette, 2021, Norgaard & Reed, 2017).

Indeed, a previous study carried out by David H. Getches in 1977 under the University of Colorado Law School advocated for a basin-scale voluntary agreement between the seven states in the CRB that was also authorized by Congress through the establishment of a new compact. Getches (1977) states that it is the “ideal” institution to have as it will more tightly incorporate localized perspectives of tribal communities, advocates, and scientists while keeping the fundamental structure of the *Law of the River* that many states are reluctant to change. This idea echoes Ostrom's logic behind Game 5, a mix of private and public authorizations over a CPR and voluntary self-enforcement of rules through collective cooperation (Ostrom, 2019). Getches reviews the current federal governing of CRB through the perspective of efficiency, equity, and sustainability with current tribal water rights status being a dominant part of equity, ultimately reaching a consensus that sharing federal authority should be a plausible approach in advancing Colorado River governance (Getches, 1977).

Furthermore, study reports by federal agencies and private organizations such as the Water and Tribe Initiative (WTI) have been delving into the interconnectedness of CRB management with tribal water rights and tribal resiliency, a previously less touched-on aspect of discussions about drought. A study published in 2021 by the WTI provides concrete details on the treaties in which the Ute Tribes, Navajo Nation, and Apache Nations were established with the United States. This highlights the fact that the majority of CRB tribes only receive federal recognition and do not have agreements with the federal government, limiting their abilities to effectively govern their water resources. This source identifies the barriers to providing access to clean water for tribes: lack of tribal agency cooperation, regulatory barriers to tribal participation, and the lack of tribal capacity to improve in-reservation water infrastructures (WTI, 2021). The Jicarilla Apache Nation gained approval from the Secretary of Interior to lease water for off-reservation use such as energy production and private industries. This marketing approach allows the Apache Nation to gain steady revenues while establishing a mutualistic relationship with the downstream users (the Bureau of Reclamation, 2018). The Colorado River Indian Tribe (CRIT), on the other hand, participated in the Pilot Conservation Program. By participating in the water conservation program, CRIT was able to closely monitor its consumptive use and water storage while receiving regular funding from the state of Arizona as well as the federal government (the Bureau of Reclamation, 2018). Additionally, graduate scholar at the University of Montana Dylan M. DesRosier tackles the need for tribal instream flow protection and ponders the potential approach of water marketing. DesRosier emphasizes forming a public consciousness toward fully realizing and accepting the sovereignty of all tribal nations (DesRosier, 2015).

Acknowledging the abundance of studies on evaluating the impacts of global climate change on CRB drought (e.g., Barnett & Pierce, 2008; Barnett & Pierce, 2009; Boehlert et al., 2015; Christensen, 2004; Edalat & Stephen, 2019), research associated with the concept of volume control (Cousins, 2016), and the general challenges facing tribal communities (e.g., Tanana et al., 2021; the Bureau of Reclamation, 2018; Stiles et al., 2015; DesRosier, 2015; Nyberg, 2015), a timely, positive, and concrete analysis on the current CRB drought-response management policies, as well as the economic benefits resulting from expanding intergovernmental water markets to quieter and more isolated tribal communities, is needed to strengthen the collective resiliency of the region. Though the idea of integrating voluntary yet widely accessible water markets to maximize the value of water and decrease supply has been established in the 20th century, empirical application of this idea is still lacking in 2021, with only Arizona adopting a state-wide banking authority and two of the ten tribes within the Partnership receiving permanent revenue from water leasing (the Bureau of Reclamation, 2018). By shedding light on this approach with reliable analysis and based on previous research, this research will evaluate the mutual benefits received from intergovernmental water markets in times of drought with the foundation of two case studies from Greece and Australia.

3. Methodology

This research will utilize a case-study approach, which will be guided by the theoretical and institutional rules elaborated in Elinor Ostrom's *Governing the Commons*, as it delves into past case studies that demonstrate the viability of governing various forms of CPRs. As I analyze the existing policies and structure of

CRB drought management, I will be able to identify their current limitations, or simply what isn't working for the Arizona farmers, tribal members of the Ten Tribes Partnership, and other appropriators who depend on the water from CRB. Then, still using *Governing the Commons* as guidance, I will examine the current tribal water rights management and evaluate the factors preventing the tribes from maximizing utilization of their tribal water claims and compelling them to lease water rights to earn much-needed revenues. As an inner case study of CRB, tribal water rights management evaluation will highlight one of the key uncertainties regarding future CRB water allocations. This will allow me to propose the approach of expanding the intergovernmental water market as a potential solution to decrease the level of future water allocation uncertainty in times of drought, increase democratic participation in CRB water management that advances voluntary collective action, and increase resilience in tribal and non-tribal communities. The last part of the research will be dedicated to analyzing Australia's Murray-Darling Basin's interstate water market and assessing the structure of the CRB intergovernmental water market with tribal communities, as well as its future mutual benefits to both sides. The purposes of this research can be summarized by the following bullet points:

- Connect CRB management policies with institutional rules detailed in *Governing the Commons*
- Identify limitations of current CRB management policies and address them using Ostrom's CPR institutional rules
- Analyze current tribal water rights in CRB and identify concrete solutions to address these limitations that are preventing the tribes from maximizing their utilization of tribal water rights
- Analyze Northern Greece's water market reform and Australia's interstate water market and identify factors of success that can be applied to the CRB intergovernmental water market
- Evaluate the future impacts of the CRB intergovernmental water market

4. Analysis of Case Studies

4.1. Chalkidiki, Northern Greece: Irrigation Water Pricing/Market Reform

Socioeconomic water shortage was triggered in Greece due to ineffective and deficient water management on the state level. Unlike the Leviathan approach taken by the United States in governing the CRB, irrigation activities in Chalkidiki can be categorized in two ways: "public collective irrigation project" and "private irrigation users" (Latinopoulos, 2005). All public irrigation activities are supervised at the local or water district level by Local Land Reclamation Boards, which are organizations controlled by and for public users (Latinopoulos, 2005). Above the level of Local Land Reclamation Boards is the General Land Reclamation Board that serves as the coordinator of the Local Boards and manages irrigation projects at the regional level (Latinopoulos, 2005). Inadequate and unregulated use of water occurred in Greece when private users, without engaging in "public collective irrigation projects," met their water demand by self-extracting from the groundwater system. In the situation of the tragedy of the commons and the prisoners' dilemma, private water users soon started to bear overwhelming costs from over-pumping the groundwater and using

more advanced technologies to extend pumping into lower water level basins. That said, a significant gap in water operational costs occurred between farmers who are private users and those involved in public collective actions when water availability dramatically declined (private users: 0.25-0.4 euro m⁻³; public users: 0.02-0.08 euro m⁻³) (Latinopoulos, 2005). This decentralized, self-enforcement system, as Ostrom (1990) proposes in Game 5, reflects its limitations: farmers overestimated the carrying capacity of the CPR, and the monitoring actions that should have been emphasized at the regional level were ineffective (Latinopoulos, 2005).

Acknowledging the need to initiate a water pricing reform system that considers the perspectives of the main stakeholder, the farmers, is crucial. While mitigating the exacerbating deficiency of water management, water markets governed through strengthened public users' organizations were considered as a solution. Because farmers in Chalkidiki are mostly family-based and subsistence, they themselves actively participated in a structured survey that examined the public's preference in supporting the operation of a water market and determined the Willingness to Pay (WTP) value for irrigation water (Latinopoulos, 2005). The value of water in the survey was based on the farmers' land values to achieve the Hedonic Property Method (HPM), which applies to evaluating the potential price of non-marketed goods like water as it is implicitly tied to the value of marketable goods like land. Echoing the idea of "reclamation," farmers were asked about socioeconomic, geographic, irrigation practices, and recent transactions and prices associated with their land.

Based on the survey, one aspect that is generally ignored by farmers is the "non-use value" components that are related to sustainability, ecosystem services, and cultural preservation (Latinopoulos, 2005). There was a lack of incentive for water conservation. Moreover, the central government prioritizes economic development over sustainability measures. This result demonstrates the need for more advanced education on environmental awareness. To address this growing concern, local environmentalists and scientists considered the combination of volumetric control while managing crop rotation that discourages high water-dependent crops to achieve a uniform water market system, yet still protecting equity among farmers.

Though water governance in Greece drastically differs from that of the United States, the process of water pricing reform can serve as inspiration to the intergovernmental water market proposal in the CRB. Just as the water price reformers took on an interactive approach to consider perspectives of the subsistence farmers through surveys on WTP, the CRB, bigger in scale, needs to also embrace the multi-perspectives of various stakeholders through direct surveying to better align the interests between producers and consumers.

4.2. Murray-Darling Basin, Australia: Interstate Water Market Established to Combat Drought

The Murray-Darling Basin (MDB) has a total reservoir storage of 28.3 MAF, the primary water use of which comes from agriculture (Wildman & Forde, 2012). The basin consists of two major rivers, the Darling (2740 km long) and the Murray (2530 km long) (Figure 2.) A water basin similar to the scale of the CRB, the Murray-Darling Basin's governance is complex, consisting of six different levels of management with a strong emphasis on state-level regulation: the central Australian government, the Australian Capital Territory, and the four states (New South Wales, Queensland, Victoria, and South Australia) that the MDB reaches (Connell & Grafton, 2011).

The 1915 River Murray Waters Agreement became the first interstate water management policy proposed by the Australian central government to allocate water among the four states. Though this agreement was established upon the agreements between states, the agreement failed in the 1980s due to increasing pressure on improving basin-wide development and salinity of the water (Connell & Grafton, 2011). The management of the MDB transitioned into the responsibility of the Murray-Darling Initiative, consisting of agencies of the central government. This reform in MDB management stimulates sustainability awareness and agriculture production, making them “institutional fold” (Connell & Grafton, 2011).

However, this Leviathan approach was soon altered by Act 38 of *Parliament of Australia* in 1993, highlighting the need for cooperation among all six governments. Immediately following this determination in 1994, intergovernmental agreements brought forward by the Council of Australian Governments (COAG) initiated more progressive basin-wide water reforms that lay the path forward for future reform regarding overallocation in times of drought. In 2003, Living Murray First Step was established by the Ministerial Council to advance inter-jurisdictional activity among different management levels to collectively address environmental degradation and recover 500 gigaliters of water (Connell & Grafton, 2011). However, this decision was not sufficient to thoroughly mitigate all ecosystem challenges facing MDB (Connell & Grafton, 2011). This led to the Australian government passing the Water Act of 2007 that centralized a great amount of authority into the central government, which marks a significant action in the past century of water management reforms (Connell & Grafton, 2011). Dominating the centralization of MDB reform is the establishment of an interstate water market, which created flexibility in meeting the water supply and demand. Though the MDB interstate water market was established in 1989, a significant increase in water trading happened in intense drought from 2005 to 2009 as water allocation was curtailed (Wildman & Forde, 2012). Utilization of the interstate water trading market was mostly between irrigators in New South Wales, whose crops are not water-dependent, and farmers in South Australia and Victoria, who need water to keep their cattle alive in times of drought (Wildman & Forde, 2012). Moreover, interstate water trading provides an important source for municipal water usage.

Interstate voluntary water trading in MDB successfully mitigated overallocation of water markets as the total withdrawals from the basin would be sufficient to satisfy the unique demand from each region based on its geographic location. Yet, the interstate voluntary water trading would ensure the region's long-run sustainable development because communication between states and various stakeholders is needed to achieve joint strategies and put the water to its most efficient use (Wildman & Forde, 2012). This form of action reflects the collective choice process that Ostrom highlights in an institutional change (Ostrom, 1990). From exercising great dependency on state governance in the past as irrigation was supervised on the district-level to realizing the mounting external yet common pressure from the change in climatic conditions and overallocation, the states themselves gradually sought more collective and centralized coordination.



Figure 2. This map of MDB portrays the geographic features of the basin, demonstrating its interstate scale and providing context for the location of the two major rivers, the Murray and the Darling. (Source: Public Domain)

5. Analysis on CRB Tribal Water Management

Under the “honest regime” of management, time must be devoted to changing rules established by a central authority. Resolving the tribal nation’s water claims by the federal government has been delayed (Figure 3, 4). This creates confusion for tribes to fully acknowledge and manage their water use. Moreover, insufficient funding and financial support prevent tribes from fully utilizing their water (the Bureau of Reclamation, 2018). The disparity of opportunities available to tribes varies greatly, which poses challenges for the Partnership to advance with collective action. Therefore, a close analysis of institutional choice on the framework of current tribal water management, taking into consideration the benefits and costs, is needed, as highlighted by the *Governing of the Commons*.

Due to the interstate and international scale of CRB, which consists of many appropriators and inconsistent fluctuations of resource unit flow, it is more costly to acquire accurate information on resource use and its actual values (Ostrom, 2019). This cause-effect relationship has impacted the tribal nations in the Partnership: lack of funding prevents the tribal nations from recognizing or managing their water claims because no monetary investments can be put into establishing a reliable monitoring system that tracks the resource condition, quality, price, and volume regularly; therefore, many tribes fail to provide timely data on their water usage. This further hinders effective “choices” to be made by the Reclamation that benefits tribal developments. Moreover, as the tribal nations along the CRB govern independently, a potential method of acquiring a monitoring system is absent. A monitoring system can take the form of an intergovernmental water master. The monitoring group, which can be voluntarily formed by cooperative actions from members of the tribal nations themselves and unrelated to the influence and authority of the states but

acknowledged by the federal government, would provide more scientifically detailed information to individual nations on their water usage of the Colorado River System. This transparent delivery of information will close up the disparity among water management of tribal nations, but more importantly will also highlight the benefits of an institutional change. Collectively powerful, a monitoring group would significantly demonstrate and strengthen the impression of governance legitimacy of tribal nations, which would attract more public awareness on the tribal's insufficient development of water claims and rights (Figure 5). A monitoring group would help resolve one of the starkest obstacles preventing progress on water leasing—tribes' inability to accurately quantify their water rights due to the slow and delayed resolution of water claims by the federal government.

Table 1. The data indicates the resolved and unresolved water claims from tribes located in the Lower Basin of CRB in acre-foot (feet) per year (AFY). This chart is adapted from the 2018 Colorado River Basin Ten Tribes Partnership Tribal Water Study (study part of the public domain). (Source: Colorado River Basin Ten Tribes Partnership Study Report)

Lower Basin Decreed Water Rights and Unresolved Claims			
Partnership Tribe	State	Decreed Diversion Right (AFY)	Unresolved Diversion Claim (AFY)
Fort Mojave Indian Tribe	Nevada	12,534	
	Arizona	103,535	
	California	16,720	
Chemehuevi Indian Tribe	California	11,340	
Colorado River Indian Tribes	Arizona	662,402	
	California	56,846	
Quechan Indian Tribe	Arizona	6,350	
	California	71,616	
Cocopah Indian Tribe	Arizona	10,847	22,928 ¹
Total		952,190	22,928

Table 2. The data indicates the resolved and unresolved water claims from tribes located in the Upper Basin of CRB in AFY. This chart is adapted from the 2018 Colorado River Basin Ten Tribes Partnership Tribal Water Study (study part of the public domain without copyrights). (Source: Colorado River Basin Ten Tribes Partnership Study Report)

Upper Basin Federal Indian Reserved Water Rights and Unresolved Claims			
Partnership Tribe	State	Settled or Adjudicated Diversion Right (AFY)	Unresolved Diversion Claim (AFY)
Ute Indian Tribe	Utah	179,315	370,370
Southern Ute Indian Tribe	Colorado	128,939 ¹	
Ute Mountain Ute Tribe	Colorado	100,184 ¹	
Jicarilla Apache Nation	New Mexico	45,682.92	
Navajo Nation	New Mexico	606,660 ²	
	Utah		314,926 ³
	Arizona (Upper Basin)		77,049 ⁴
Total		1,060,780.92	762,345

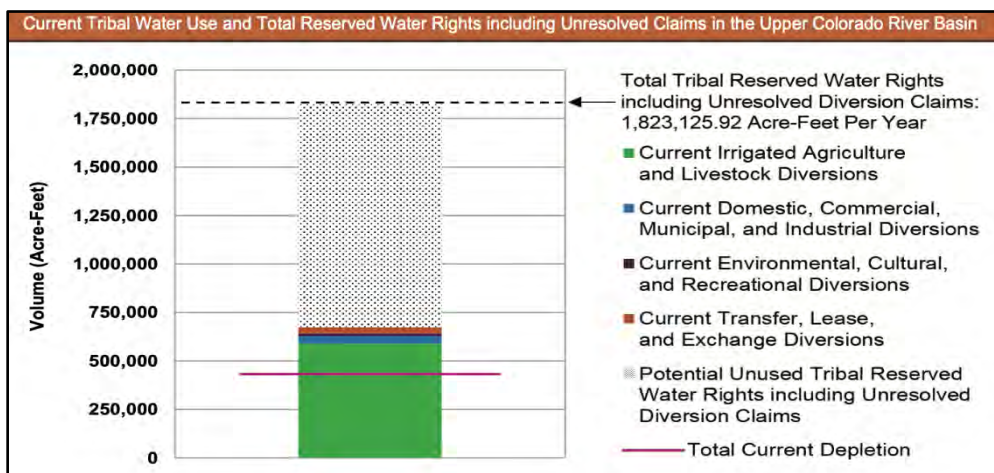


Figure 3. The graph indicates the utilized volume of tribal water rights versus the non-utilized, or water rights that have not been resolved. Less than half of the authorized water has been used by the tribal nations in the Partnership. This graph is adapted from the 2018 Colorado River Basin Ten Tribes Partnership Tribal Water Study (study part of the public domain without copyrights). (Source: Colorado River Basin Ten Tribes Partnership Study Report)

5.1 Intergovernmental Water Market Proposal

To bring a permanent source of funding to tribal communities and achieve efficiency in water allocations, an intergovernmental water market that acknowledges the sovereign entities of tribal alliances like the Ten Tribes Partnership is an optimal solution. In 2018, 375,000 acre-feet of water were traded for \$138 million in Arizona (Schwabe et al., 2020). Determining the guidelines of water pricing between the Arizona farmers that have been impacted by the CAP decrease of water allocation and the Ten Tribes Partnership would require collective participation in surveys and monitoring between all stakeholders. This requires the establishment of a basin-wide water master governmental market council among the Colorado River tribal countries that consists of tribal water representatives proportional to the population of each federally recognized tribal nation and a federal department that is devoted to monitoring the leasing and transaction of water delivery from a tribal nation to off-reservation water appropriators. As Ostrom's (2019) institutional rules of accountability highlight, having this water market that combines the authority of the federal government and the tribal nations will grant substantial autonomy to the tribal nations that decreases the cost of transformation, an advantage already reflected in the Murray-Darling Basin interstate water market. More specifically, this water market will directly decrease the amount of time and financial resources spent on convincing the central authority to acknowledge the transaction request and waiting for an answer.

The purpose of this federally recognized council is to allow transparency in water supply, volume control, and market price among different tribal nations; however, it does not establish a unifying market value to water due to the necessity of acknowledging the regional differences among countries and states. Price per acre-

foot water will fluctuate with the natural tendency of a free market and is designed to respond sensitively to the impacts of exacerbating global climate change and the more frequent occurrence of environmental challenges in communities. Instead of only receiving survey input from the buyers on WTP, as indicated by a case study in Northern Greece, which could result in unreliable and skewed results in an underestimation of water price and ignored sustainability awareness, tribal communities' input would ensure their willingness to supply (WTS) in consideration of cultural and environmental conservation. This voluntary intergovernmental water market system would serve as a unifying factor to ensure the incorporation of tribal perspectives in water negotiation but also addresses one of the challenges faced by the tribal nations in water leasing: the lack of a reliable market (Figure 6).

Leasing water in the form of free enterprise where the factors of production lie in the hands of individual appropriators (farmers, tribal nations, etc.) requires a fundamental change in the current managing system of the Colorado River with fierce governmental interventions. This change in the management system will limit restrictions imposed by the Nonintercourse Act, which implicitly prevents tribal nations from leasing or selling water rights without authorization from Congress (Wildman & Forde, 2012; Nyberg, 2015). As argued by Nyberg (2015), tribal nations leasing their water rights for economic benefits and putting scarce resources to more efficient use is merely a case of "use it or lose it" (pp. 198). Some policymakers may argue that as tribal nations lease their water to non-reservation users, urban developments will expand that further isolate the Native American communities. However, an intergovernmental water market is solely need-based under free enterprise. Such political tensions and legal obstacles do not diminish the potential of establishing the intergovernmental water market, as it still serves as a solution to the exacerbating socioeconomic drought in the CRB.

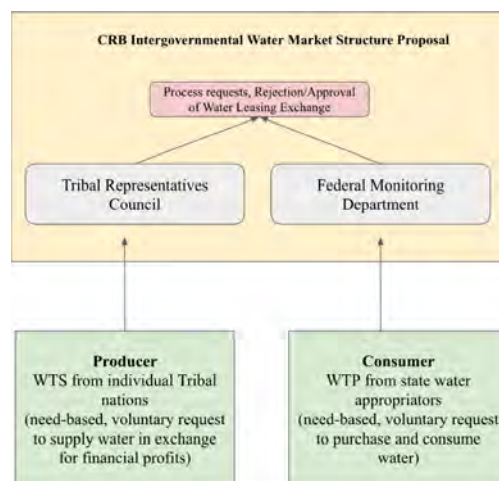


Figure 4. This diagram elaborates on the ideal structure of the CRB intergovernmental water market. The CRB Intergovernmental Water Market is a need-based-driven market. "Producer" refers to individual tribal nations while "Consumer" refers to non-reservation water users. Based on WTS and WTP, both sides of the stakeholder submit requests to designated councils and departments within the market for collective review.

6. Discussion

This study details the proposal of establishing a Colorado River basin-wide intergovernmental water market and acknowledges the need for an institutional change to the *Law of Rivers*. A complete scope of review on the legal complexities of carrying out an institutional change is beyond the scope of this paper and remains a potential area of study in future research that tackles the obstacles faced by each tribal nation in leasing water through in-field studies and thorough interviews. Moreover, available resources are limited to fulfill field observations that would capture a greater depth of intricacy in tribal responses to drought, as available and published information in former studies and databases is greatly constrained. Moreover, potential non-binding and implied agreements are in place between tribal nations that are not recorded in digital documents, which is open for future research opportunities. This limitation is due to the unique water management policies and challenges each tribal nation has due to unique geographic locations and cultures. Complexity must be acknowledged in analyzing the limitations of tribal water rights. The solution of establishing an intergovernmental market to provide a stable stream of funding might not apply to specific nations due to existing policies hindering the leasing of tribal water. This uncertainty opens areas for future research to delve into more targeted and concrete institutional change for tribal nations. These uncertainties can be collectively addressed by having a unified water market that promotes exchange in water.

As the socioeconomic drought in the western United States increasingly gains national attention, it is crucial for this study to acknowledge progressive efforts have been made on the federal level. According to United States Representative Anna Eshoo, the bipartisan infrastructure bill that was passed by the Senate in July of 2021 invests \$8.3 billion in drought mitigation in western states that include water storage, recycling, conservation, and efficient projects. More specifically, a \$3.5 billion funding is dedicated to tribal nations for water infrastructure projects without new water settlements. This critical funding allows tribal nations from the CRB to strengthen their water delivery systems while leaving the inevitable water claim conflicts between tribal nations and the non-reservation stakeholders unresolved. A thorough reflection on the governance structure of the CRB and an innovative approach of regulating water allocation to maximize efficiency in scarce resource use is still needed to address the fundamental threats of drought.

7. Conclusion

As the federal government creates drought contingency plans to encourage conservation of water and decreases allocation of water for junior water rights holders, more localized, timely, and direct actions among the appropriators are crucial to put water, an increasingly scarce resource, to most efficient use. Through the analysis of Ostrom's institutional rules (2019) on CPR's governance, collective voluntary actions are needed to seek a balance between centralized governmental control and privatization, thereby mitigating the tragedy of the commons and the prisoner's dilemma. By gathering factors of success from water market reform in Northern Greece and the establishment of an interstate water market in the Murray-Darling Basin in Australia, an intergovernmental market proves to be an accessible

solution. This proposal would entail the appointment of a water master organization among tribal nations that collectively monitors the consumption, storage, and transaction of water while providing accountability and transparency to tribal water claims. Moreover, by forming this need-based water market that meets the demand for revenue and the demand for water from tribal nations and state stakeholders respectively, individual stakeholders and water users will collectively approach understanding the value of water in a more holistic manner.

References

- Barnett, T. P., & Pierce, D. W. (2008). When will Lake Mead go dry?. *Water Resources Research*, 44(3), n/a. doi:10.1029/2007WR006704
- Barnett, T. P., & Pierce, D. W. (2009). Sustainable water deliveries from the Colorado River in a changing climate. *Proceedings of the National Academy of Sciences of the United States of America*, 106(18), 7334–7338. <https://doi.org/10.1073/pnas.0812762106>
- Boehlert, B., Fitzgerald, E., Neumann, J. E., Strzepek, K. M., & Martinich, J. (2015). Effects of Greenhouse Gas Mitigation on Drought Impacts in the United States. *Weather, Climate & Society*, 7(3), 255–272. <https://doi.org/10.1175/WCAS-D-14-00020.1>
- Borunda, A. (2021, May 7). 'Megadrought' persists in western U.S., as another extremely dry year develops. *National Geographic*.
- Boström, M., & Davidson, D. J. (2018). *Environment and Society: Concepts and Challenges*. Springer International.
- Bureau of Reclamation, L. C. R. (2018, December). Colorado River Basin Ten Tribes Partnership Tribal Water Study. Tribal Water Study Report. <https://www.usbr.gov/lc/region/programs/crbstudy/tws/finalreport.html>.
- Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C., and Famiglietti, J. S. (2014). Groundwater depletion during drought threatens future water security of the Colorado River Basin, *Geophys. Res. Lett.*, 41, 5904– 5911, doi:10.1002/2014GL061055.
- Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P., & Palmer, R. N. (2004, January). The effects of climate change on the hydrology and water resources of the Colorado River Basin. Retrieved July 27, 2021, from <https://doi.org/10.1023/B:CLIM.0000013684.13621.1f>
- Congressional Research Service. (2021, July 15). Management of the Colorado River: Water (C. V. Stern & P. A. Sheikh, Authors; Report No. R45546). Congressional Research Service. <https://crsreports.congress.gov/product/pdf/R/R45546>
- Cousins, J. J. (2017). Volume control: Stormwater and the politics of urban metabolism. *Geoforum*, 85, 368-380. <https://doi.org/10.1016/j.geoforum.2016.09.020>
- Davenport, J. (Spring, 2007). Softening the Divides: The Seven Colorado River Basin States' Recommendation to the Secretary of the Interior Regarding Lower Basin Shortage Guidelines and the Operation of Lakes Mead and Powell in Low Reservoir Conditions. *University of Denver Water Law Review*, 10, 287. <https://advance.lexis.com/api/document?collection=analytical-materials&id=urn:contentItem:4VW5-7SCO-00SW-508G-00000-00&context=1516831>.

- Dawadi, S., & Ahmad, S. (2012). Changing climatic conditions in the Colorado River Basin: Implications for water resources management. *Journal of Hydrology*, 430–431, 127–141. doi:10.1016/J.JHYDROL.2012.02.010
- DesRosier, Dylan M. (2015). "Tribal Water Rights Settlements and Instream Flow Protection." Graduate Student Theses, Dissertations, & Professional Papers. 4433. <https://scholarworks.umt.edu/etd/4433>
- Edalat, M. M., & Stephen, H. (2019). Socio-economic drought assessment in Lake Mead, USA, based on a multivariate standardized water-scarcity index. *Hydrological Sciences Journal/Journal Des Sciences Hydrologiques*, 64(5), 555–569. <https://doi.org/10.1080/02626667.2019.1593988>
- Getches, D. (Summer, 1997). Article: Colorado River Governance: Sharing Federal Authority as an Incentive to Create a New Institution. *University of Colorado Law Review*, 68, 573. <https://advance.lexis.com/api/document?collection=analytical-materials&id=urn:contentItem:3S3T-T610-00CV-N03B-00000-00&context=1516831>.
- Latinopoulos, P. (2005). VALUATION AND PRICING OF IRRIGATION WATER: AN ANALYSIS IN GREEK AGRICULTURAL AREAS.
- Leblanc, M. J., Tregoning, P., Ramillien, G., Tweed, S. O., and Fakes, A. (2009). Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia, *Water Resour. Res.*, 45, W04408, doi:10.1029/2008WR007333.
- McNabb, D. E. (n.d.). The State of America's Water Resource - 1. In D. E. McNabb, *Water Resource Management*. Springer International Publishing. https://doi.org/10.1007/978-3-319-54816-6_1 (Original work published 2017)
- Monitoring Strategy and Assessment Workgroup, U. S. E. P. A. (2018). Developing a Tribal Water Monitoring Strategy. Supplement to the Clean Water Act Section 106 Tribal Guidance. EPA. <https://www.epa.gov/water-pollution-control-section-106-grants/developing-tribal-water-monitoring-strategy-supplement>.
- Mount, J., Hanak, E., Chappelle, C., Colby, B., Frank, R., Gartrell, G., Gray, B., Kennedy, D., Lund, J., & Moyle, P. (n.d.). Improving the federal response to western drought. *California Journal of Politics and Policy*. <https://doi.org/10.5070/P2cjpp8331897>
- Nyberg, J. (2015). The Promise of Indian Water Leasing: An Examination of One Tribe's Success at Brokering Its Surplus Water Rights, 55. *Nat. Resources J.* 181. Available at: <https://digitalrepository.unm.edu/nrj/vol55/iss1/8>
- Nobel, J. (2016, September 16). These are the forgotten victims of the west's drought. *National Geographic*. Retrieved July 24, 2021, from <https://www.nationalgeographic.com/science/article/nevada-drought>
- Norgaard, K.M., Reed, R. Emotional impacts of environmental decline: What can Native cosmologies teach sociology about emotions and environmental justice?. *Theor Soc* 46, 463–495 (2017). <https://doi.org/10.1007/s11186-017-9302-6>
- Ostrom, E. (2015). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press. <http://doi.org/10.1017/CBO9781316423936>
- Schwabe, K., Nemati, M., Landry, C., & Zimmerman, G. (2020). Water markets in the western united states: Trends and opportunities. *Water (Basel)*, 12(1), 233. <https://doi.org/10.3390/w12010233>

- Spivak, D. (2021). The Colorado River Drought Contingency Plan: An Opportunity for Exploring Demand Management Through Integrated and Collaborative Water Planning. *Natural Resources Journal*, 61(2), 173-204. Retrieved July 28, 2021, from <https://www.jstor.org/stable/27034470>
- Tanana, H., Garcia, J., Olaya, A., Colwyn, C., Larsen, H., Williams, R., & King, J. (2021, April). Universal access to clean Water for Tribes in the Colorado River Basin. Water & Tribes Initiative.
- Todd Fitchette. (April 9, 2021). Drought hitting home in California, Arizona. The Corn & Soybean Digest. <https://advance.lexis.com/api/document?collection=news&id=urn:contentItem:62D8-56D1-JBKS-P1XW-00000-00&context=1516831>.
- Stiles, C., Umphlett, N., Rattling Leaf, J., Shulski, M.D., Kluck, D., Hayes, M., & McNutt, C. (2015). Improving Climate Resiliency in Tribal Communities: Partnering for Change in the Missouri River Basin. *Water Resources Impact*, 17(4), 15-17. Retrieved July 28, 2021, from <https://www.jstor.org/stable/wateresoimpa.17.4.0015>
- Ward, F., & Michelsen, A. (2002). The economic value of water in agriculture: concepts and policy applications *S. Water Policy*, 4, 423-446.
- Wildman, R. A., Jr., & Forde, N. A. (2012). Management of Water Shortage in the Colorado River Basin: Evaluating Current Policy and the Viability of Interstate Water Trading 1 . *JAWRA Journal of the American Water Resources Association*, 48(3), 411-422. doi:10.1111/J.1752-1688.2012.00665.X



Comparing the Response Rate, Remission Rate, and Effective Age Range of High-Frequency Left Transcranial Magnetic Stimulation (HFL rTMS) and Low-Frequency Right Transcranial Magnetic Stimulation (LFR rTMS) in Treatment of Major Depressive Disorder (MDD)

Xintong Chen

Author Background: *Xintong Chen grew up in China and currently attends York School in Monterey, California in the United States. Her Pioneer research concentration was in the field of psychology and titled “Clinical Psychology.”*

Abstract

Repetitive transcranial magnetic stimulation (rTMS) is a type of non-invasive brain-stimulation technique that has been used clinically to treat patients with major depressive disorder (MDD). It is especially beneficial for patients with treatment-resistant depression (TRD) who fail to respond to traditional treatment with medication. While the conventional application of rTMS involves high-frequency stimulation targeted at patient’s left dorsolateral prefrontal cortex (DLPFC), a variation that applies low-frequency stimulation on the right DLPFC has been suggested by previous studies to have comparable efficacy in the treatment of MDD. The relatively novel low-frequency right (LFR) rTMS has several advantages over high-frequency left (HFL) rTMS in that it can significantly decrease the cost, increase public accessibility, and reduce the side effects during and after the treatment. This paper argues that LFR rTMS can be a potential alternative to HFL rTMS with more evidence gathered. Currently, most experimental trials are designed to test the immediate treatment response of HFL rTMS and LFR rTMS, while data on the remission rates and long-term response rates are limited. In addition, the role that patients’ ages plays in treatment response remains controversial. Therefore, an experiment is designed in this paper to test the remission rates, long-term response rates, and age as a factor in the treatment outcomes of HFL rTMS and LFR rTMS. If the results prove that the two variations have comparable remission rates, long-term effect, and age range responding to the treatment, the study provides more evidence for LFR rTMS to replace the conventional HFL rTMS treatment.

Introduction

Repetitive Transcranial Magnetic Stimulation (rTMS) is a type of brain stimulation that has been proven to have antidepressant effects on patients with Major Depressive Disorder (MDD). While the high-frequency left (HFL) rTMS is the most widely used type of rTMS with a confirmed efficacy, a growing number of studies suggest that low-frequency right (LFR) rTMS can be a potential alternative to the conventional HFL rTMS as it demonstrates comparable efficacy in treating MDD. Further research is still needed to compare the two variations. Currently, data on the remission rate, long-term antidepressant, and influence of age on the clinical outcome for LFR rTMS is still limited. This study is designed to explore these three variables in order to provide a more comprehensive picture of LFR rTMS.

What is Repetitive Transcranial Magnetic Stimulation (rTMS)?

With an annual prevalence rate as high as 5% - 15%, Major Depressive Disorder (MDD) is one of the most widespread mental disorders around the world. MDD causes personal distress for patients as it is associated with social and physical dysfunction and a high suicide rate. Symptoms of major disorder episodes include lack of pleasure, loss of interest in activities, insomnia, and appetite and weight disturbance, etc. (Kayser et al., 2012). While the traditional methods of intervention include antidepressant medications, not all patients respond to such treatments. For patients developing drug-resistance tendencies in their treatments of MDD, repetitive magnetic stimulation (rTMS), a novel, non-pharmacological technique of brain stimulation, is proposed to serve as a treatment method for treatment-resistant depression (TRD) (Peng et al., 2018).

The application of rTMS is based on Faraday's law concerning the relationship between electricity and magnetic fields. In a conventional high-frequency left-sided (HFL) rTMS treatment, a coil circuit is placed on a specific location above the patient's scalp. When electricity is being conducted into the circuit of the apparatus, a magnetic field is generated around the area of stimulation and thus induces a small amount of electrical current to stimulate the left dorsolateral prefrontal cortex (DLPFC). When the stimulation is applied repetitively with a frequency greater than 5 Hz, patients often report a superior antidepressant effect compared to the sham control group. HFL rTMS was approved by the United States Food and Drug Administration (FDA) in 2008 as a treatment for MDD (Peng et al., 2018).

rTMS has several advantages over traditional treatments of MDD other than its efficacy in treating TRD. First, antidepressant medication is not applicable for some patients with other medical conditions. Patients with myocardial infarction (MI), for example, experience various side effects using tricyclics and monoamine oxidase inhibitors (MAOIs). The use of selective serotonin reuptake inhibitors (SSRIs) is also limited as they interact with cytochrome P450, a component in some cardiac drugs. Thus, rTMS can be a possible alternative under this circumstance due to its non-pharmacological nature (Valiengo et al., 2013). Second, compared to another type of brain stimulation, electroconvulsive therapy (ECT), rTMS causes little cognitive

impairment or other side effects. In neuropsychological evaluation tests from a comparison study conducted by Rosa et al. (2006), patients who received ECT showed significantly decreased visual memory and aspects of executive functioning (verbal fluency) compared to the rTMS group. Therefore, rTMS is a meaningful clinical method to develop in addition to traditional treatments.

There are several variations of the original HFL rTMS, such as low-frequency right right-sided (LFR) rTMS, intermittent theta-burst stimulation (iTBS), deep TMS, bilateral stimulation, and priming stimulation. All those methods have proven potentially effective in treating specific mental disorders, but further research is required to confirm their practicality (Fitzgerald & Daskalakis, 2011). In a meta-analysis examining results from 81 studies, only five types of variations are proven to be more significantly effective than the sham group: priming low-frequency, bilateral, theta-burst stimulation, high-frequency, and low-frequency (Brunoni et al., 2017). One of the variations, LFR rTMS, involves stimulating the right DLPFC with frequency less than or equal to 1 Hz. It is suggested that LFR rTMS can be a potential alternative to the conventional HFL rTMS as it has been shown by studies to have an equal response rate and higher public accessibility (Miron et al., 2020a).

Theories of Mechanism

The mechanism for the antidepressant effect of rTMS remains unclear. However, there are numerous theories that can potentially explain the antidepressant effect. Luan et al. (2020) reviewed evidence from preclinical studies and summarized the mechanisms of rTMS. Luan et al. proposed that rTMS functions by inducing anti-inflammatory effects and anti-oxidative stress effects, enhancing synaptic plasticity, altering levels of monoamine neurotransmitters, or changing the activity of the hypothalamic-pituitary-adrenocortical axis. All mechanisms are supported by a portion of prior studies, yet other studies have yielded conflicting results.

Among all theories of mechanisms, the most accepted view is that rTMS causes hippocampal neurogenesis and alters synaptic plasticity through long-term potentiation and long-term depression. Hippocampal neurogenesis refers to the creation of new neurons in the hippocampal region, which can affect or be affected by brain areas related to mood and cognition. Hippocampal neurogenesis is often an effect of antidepressant medications, and thus has a potential causal relationship to successful MDD treatments (Barlow & Targum, 2007).

Synaptic plasticity refers to the phenomena that the strength of connectivity between synapses can be altered and is associated with cortical excitability of a specific brain region. Cortical excitability is commonly measured by resting motor threshold (RMT) or motor evoked potential (MEP). RMT refers to the minimum amount of stimulation intensity required to induce an observable twitch in certain muscles. MEP is defined as the average regional response to stimulations measured through electromyography. The effect on synaptic plasticity is where HFL rTMS differs from LFR rTMS in mechanism. Fitzgerald, Fountain, and Daskalakis (2006) reviewed 10 studies of high-frequency rTMS and 19 studies of low-frequency rTMS examining their post-train effects on cortical excitability. They concluded that in rTMS with frequency greater than 5 Hz, the magnetic stimulation demonstrates enhanced neuronal excitability with

an increase in MEP size. On the contrary, the LFR rTMS with a frequency less than 1 inhibits neuronal excitability with a reduction in MEP size.

The fact that altering cortical excitability can be utilized therapeutically in treatments of MDD is due to the abnormal hypoactivity or hyperactivity of the DLPFC region within patients with depression. To discover that hypoactivity of left DLPFC is related to depression, Bexter (1989) studied cerebral glucose metabolism through positron emission tomography (PET) in 30 patients with three types of disorders: 10 with unipolar depression, 10 with bipolar depression, and 10 with obsessive-compulsive disorder (OCD). Mean glucose metabolic rates for the left DLPFC were collected and divided by the rate of the ipsilateral hemisphere. The final measures of cortical excitability for patients with depression (a mean glucose metabolic rate of 1.05 ± 0.05 for unipolar depression and 1.04 ± 0.05 for bipolar depression) were consistently lower than those in the control group (mean glucose metabolic rate = 1.12 ± 0.06) or those with OCD (mean glucose metabolic rate = 1.15 ± 0.05), thus proving the hypoactivity. In contrast, a study done by Cao et al. (2018) on the right DLPFC has indicated the opposite phenomena. Glucose and oxygen consumption is at a higher level for patients with depression, proving the hyperactivity of the region. Therefore, HFL rTMS influences the hypoactivity of left DLPFC by elevating cortical excitability, while LFR rTMS counteracts the abnormal right DLPFC hyperactivity by decreasing cortical excitability.

Efficacy of HFL rTMS and LFR rTMS

HFL rTMS

The antidepressant effect of HFL rTMS has been supported by a large number of controlled studies and is approved by the FDA as a treatment of MDD. In a meta-analysis of the efficacy of HFL rTMS, 1164 drug-resistant patients with a diagnosis of depression were compared. Among the total sample, 558 patients received the HFL rTMS treatment, while 558 received the sham treatment. The weighted mean effect size was 0.39, which indicates a clear antidepressant effect. In addition, an analysis of variance (ANOVA) test in the same study demonstrated that there is little difference in efficacy between patients receiving HFL rTMS with or without medication resistance. Therefore, drug-resistance is not a factor influencing the outcome of rTMS treatments. Side effects are well tolerated by patients and no observed seizures or deaths in the treatment group have been reported. The most common forms of side effects include headaches, dizziness, or nausea, which are milder compared to those of other types of brain stimulations (Schutter, 2009). Furthermore, while most literature assessing the efficacy of HFL rTMS focus on patients that receive no more than one treatment session per day, Chen, Hudaib, Hoy, and Fitzgerald (2020) systematically reviewed the efficacy of so-called accelerated rTMS schedules, in which patients were given HFL rTMS more than once daily. 12 records were analyzed with reported response rates ranging from 36% to 56%, not significantly different from the standard once-daily rTMS treatments.

The long-term effect of HFL rTMS treatment is not fully explored and the related literature remains scarce. One double-blind, randomized controlled trial investigated this aspect of HFL rTMS (pulse frequency = 10 Hz, intensity = 120%

of the RMT). 301 patients enrolled in the study were randomly assigned to the active rTMS group or sham group and received the treatment for 6 weeks. Patients with a 25% decrease in the Hamilton Rating Scale for Depression (HAM-D) score proceeded to a transitional period with gradual reduction of rTMS sessions and eventually shifted to a maintenance monotherapy with medication. After 24 weeks, a follow-up clinical assessment was conducted to evaluate the durability of rTMS. The results showed that among all patients who responded to active HFL rTMS treatment, 10 of 99 (10%) relapsed, 38 of them (38.4%) met the criteria for symptom worsening according to Clinical Global Impression Severity of Illness (CGI-S), and 32 of 38 (84.2%) patients with worsened symptoms benefitted again through the reintroduction of rTMS. The researcher compared the results with previous meta-analysis about the relapse rate of antidepressant medications and ECT, concluding that rTMS yields more favorable results in terms of long-term clinical outcome (Janicak et al., 2010).

However, another meta-analysis by Kedzior et al. (2015) examining 16 randomized controlled trials drew the opposite conclusion. Patients (N = 495) included in the analysis received 5-15 HFL rTMS sessions with no maintenance treatments. The antidepressant effect was shown to be statistically significant yet small in magnitude, unlike what is mentioned by the study above. The researcher suggested that a better antidepressant effect might be achieved when more sessions of rTMS are performed. Consensus has not been reached yet over the issue of whether HFL rTMS has long-term efficacy in the treatment of MDD.

LFR rTMS

Compared to the evidence base of HFL rTMS, less data was collected to test the efficacy of LFR rTMS. According to an evidence-based guideline updated in 2020 by Lefaucheur et al., HFL rTMS has been rated as having Level A efficacy (definite efficacy), while LFR rTMS was rated Level B efficacy (probable efficacy). However, whether LFR rTMS is less effective in treatment of MDD is constantly debated, and a growing amount of evidence suggests that LFR rTMS has comparable response and remission rates compared to HFL rTMS. For example, Cao, Deng, Su, and Guo (2020) reviewed 12 studies in a meta-analysis and suggested equivocal efficacy between HFL rTMS and LFR rTMS. In the analysis, a total of 361 patients were included with 175 receiving HFL rTMS and 186 receiving LFR rTMS. The response rates were analyzed with an odds ratio of 1.08 (95% confidence interval = 0.88-1.34), while the remission rates had a similar odds ratio of 1.29 (95% confidence interval = 0.54-3.10).

Following a similar procedure, Berlow, Zandvakili and Philip (2020) reviewed 12 randomized controlled trials with a total number of 627 patients (320 were treated with LFR rTMS, while 307 were treated with HFL rTMS). The response criterion was a 50% decrease in depression rating scales. The results show that 47% (292 out of 627) of subjects responded to the treatment overall, with a 45% response in the LFR rTMS group and a 48% response in HFL rTMS (odds ratio = 1.12; 90% CI = 0.85-1.48). Remission rates also indicate no significant differences between the two groups, with a 26% rate for LFR rTMS group and 25% for the HFL rTMS group (odds ratio = 0.92; 90% CI = 0.64-1.33). Therefore, no significant difference in efficacy was detected between the two types of rTMS treatments in terms of clinical response and remission.

Limitations of current literature

Although the arguments made by meta-analyses comparing the HFL rTMS and LFR rTMS offer great insight into a potential equivalence of antidepressant efficacy, there are three primary limitations of previous literature that might undermine the validity of the conclusion: a lack of reported data on remission rates, a lack of reported data on the durability of rTMS effects, and difficulties related to the maintenance of blinding.

First, not all randomized controlled trials include data for the remission rates. Among the 12 studies examined by Cao et al. (2018), only 5 of them reported the remission rate. While the original sample size used to calculate the response rates was 361, the number of subjects was reduced to 131 when the remission rates were calculated. Similarly, only 6 studies among the 12 studies in the review from Berlow et al. (2020) recorded the remission rates. As a result, 627 subjects were analyzed for response rates, whereas only 431 patients were involved in the calculation of remission rates. Therefore, the data of remission rates may not be as valid as that of response rates due to a significant reduction in sample size.

Second, few studies have been conducted to measure the long-term antidepressant effects of either HFL rTMS and LFR rTMS. In both of the two meta-analysis studies by Cao et al. (2018) and Berlow et al. (2020), no randomized controlled trials collected data to assess the severity of patients' depression after an extended period of time. Although the study conducted by Janicak et al. (2010) explored the lasting effects of HFL rTMS, the results may not be representative of the actual effect of rTMS because patients were given antidepressant medications after the rTMS treatments. Further research on the durability of HFL rTMS and LFR rTMS is needed to have a full understanding of their therapeutic value.

Lastly, the procedural blinding is found to be challenging in rTMS trials. Mogg et al. (2008) suggest that according to their research, only a few of the previous studies conducted patient blinding successfully, and none of them achieved rater blinding. In their own controlled experiment regarding the antidepressant effect of HFL rTMS, an overall 67% (34 out of 51) of subjects guessed the group that they belong to. 70% of patients in the active treatment group believed that they were receiving real treatment, while only 38% in the sham group did. The gap between the percentages indicated a failed attempt at subject blinding, which may have influenced the clinical outcomes. The same research team also summarized and described the options for commonly used sham coils. They suggest that most studies used a tilted, real TMS coil to both ensure a scalp sensation and avoid stimulating targeted areas. However, this type of practice might also alter the cortical activity through stimulation. Mogg et al. (2008) used a sham coil designed to be incapable of stimulation, yet the lack of scalp sensation compromised the effect of subject blinding. Therefore, designing or choosing a proper sham coil can be challenging while also important, as it is closely associated with the successful blinding of the experiment.

Patient-related Factors Influencing the Clinical Outcome

Gender

Although not yet well researched, there is evidence suggesting that female patients may have better treatment outcomes after rTMS sessions compared to male patients. Kedzior, Azorina, and Reitz (2014) conducted a meta-analysis on 54 published sham-controlled studies and discovered that higher rates of treatment response often occurred in studies with a higher proportion (> 60%) of female participants. The data analyzed in this study include both patients receiving HFL and LFR rTMS, so it is indicated that gender plays a role in both variations of rTMS treatment.

Age

There is also evidence that age is a predictor of clinical outcomes of rTMS. For HFL rTMS, the treatment response appears to be negatively correlated to age. One study reviewed data from 42 clinical sites and compared the treatment response of 307 outpatients, who were divided into two age groups. The group with patients younger than 55 years old was considered the younger group, and the other with the age range beyond 55 was considered the older group. Most of them received HFL rTMS as their treatment. The final results testing the clinical outcomes indicate that the younger patients tend to have a more favorable antidepressant effect (Carpenter et al., 2012).

A similar conclusion was also drawn on LFR rTMS. Pallanti et al. (2012) analyzed data from 102 treatment-resistant patients who received LFR rTMS, assessing the influence of age on the efficacy. They defined subjects younger than 60 as the younger group and those older than 60 as the older group. After three weeks of stimulation sessions, they discovered a significant negative correlation between age and symptom reduction (as tested by HAMD scores). Thus, LFR rTMS seems to be less effective for patients with higher age.

However, whether age can be a predictor of clinical outcome for rTMS is still highly controversial. The study conducted by Sabesan et al. (2015) specifically reviewed research on the efficacy of rTMS on elder patients. They concluded that age itself is not a predictor of outcome since the elders also showed outcomes as favorable as those of young patients. The response rate (>50% reduction in HAMD score) in the elder population reached as high as 58.5%. In one controlled study, the elder group had in fact demonstrated a better response than the younger group. Therefore, further research needs to be conducted to address and explain the inconsistent outcomes.

Potential Advantages of LFR rTMS over HFL rTMS

The reason for researching LFR rTMS as a potential alternative to HFL rTMS is that LFR rTMS can decrease the number of side effects, reduce the cost of rTMS treatments and increase public accessibility. In past studies, patients have shown higher tolerability toward the LFR-rTMS and reported fewer sensations of pain after the treatment. The study conducted by Kaur et al. (2019) asked the patients to rate the overall sensations, pain, muscle twitching, and discomfort after

receiving both HFL rTMS (10 Hz at 100% RMT) and LFR rTMS (1 Hz at 100% RMT), and the results showed that patients rated lower in all those measures for LFR rTMS. There is also evidence that patients receiving LFR rTMS have lower rates of seizure compared to those receiving HFL rTMS (Miron et al., 2020a).

Another advantage of LFR rTMS is that the apparatus needed for the treatment is much less expensive compared to HFL rTMS. While the conventional HFL treatment requires the use of a cooled figure of 8 (Fo8) coil, Miron et al. developed a non-cooled and non-focal parabolic coil to deliver LFR rTMS at a more affordable cost (Miron et al., 2020a). The effectiveness of the parabolic coil was tested by the same team, and 43 patients with treatment-resistant depression reported a 37.2% response rate and 23.3% remission rate (Miron et al., 2020b). Although the sample size is relatively small and more research is needed, it is a good indication that LFR rTMS can be a more economical option for patients with TRD.

Lastly, because LFR rTMS requires less precision in localization, the application of the stimulation coil is easier and requires less professional knowledge. The simplicity of the treatment procedure leads some researchers to suggest that LFR rTMS can be possibly used by patients themselves at home, which is especially beneficial for patients during the COVID pandemic as social interactions are limited and traveling is not recommended (Miron et al., 2020a; Caulfield & George, 2020).

Hypothesis

I hypothesize that both the HFL rTMS group and LFR rTMS group will have a better antidepressant outcome compared to the sham group in response rate, remission rate, and the long-term effect. Among two groups divided by age differences, younger patients will respond better to both HFL rTMS and LFR rTMS treatment compared to the older age group. Both age groups receiving active treatment will obtain better clinical outcomes compared to the sham group. Older patients will respond better to LFR rTMS compared to HFL rTMS.

Method

Participants

A total of 150 patients will participate in the experiment with referral of their clinicians. The age range of the patients will be controlled to be between 20 to 60, with about half of the patients ranging from age 20 to 40 and half of the patients from 40 to 60. All patients will have a diagnosis of major depressive disorder (MDD) according to the DSM-V and score higher than 7 on the Hamilton Depression Rating Scale (HAMD). Patients will not be on any types of antidepressants or attending other brain stimulation treatments such as ECT one month prior to their first session of the rTMS treatment. Left-handed patients will be excluded from the sampling because previous research suggested that a small portion of left-handed patients respond to HF rTMS better than LF rTMS (Fitzgerald, Hoy, & Daskalakis, 2020).

Materials and Apparatus

The severity of the MDD symptoms will be assessed by the 17-item Hamilton Rating Scale for Depression (HAMD-17) score. HAMD is a valid, widely used depression scale to quantify the severity levels of MDD. Leucht et al. (2013) assessed its validity through data collected from 43 studies with 7131 patients and discovered a strong correlation between the net HAMD score and the Clinical Global Impression Scale (CGI), a 7-point scale for therapists to rate the severity of patients' disorders. To meet the criteria for "treatment response" in this experiment, patients must have at least a 50% decrease in HAMD score at the end of the last rTMS session. Subjects will be considered in remission if their HAMD scores are lower than 7 at the end of the treatment. Patients will meet criteria for "long-term symptom reduction" if there is still a 40% decrease in the HAMD score 6 months after the end of the treatment.

The apparatus used in the study will be a Medtronic Magpro R30 magnetic stimulator with a figure of 8 coil. It is a magnetic stimulator developed by MagVenture for research and therapeutic purposes and can run as many as 20,000 pulses in one session, which is sufficient for this experiment ("Magpro r30," n.d.).

Procedures

The study will follow a double-blind procedure. The researcher is aware of the group assignment, but clinicians giving assessments according to HAMD will not be informed. Participants will be randomly assigned to three groups through a computer program with each group consisting of 25 patients between ages 20 and 50 and 25 patients between ages 50 to 80. This division of age references the study by Carpenter et al. (2012), in which the researchers considered patients older than 55 the older group and those younger than 55 the younger group. The three groups will each receive HFL rTMS (frequency = 5 Hz) treatment, LFR rTMS (frequency = 1 Hz), and sham treatment.

The rTMS setup and parameters in this study will be similar to the randomized controlled trial designed by Fitzgerald et. al in 2009 with some modifications since Fitzgerald's trial has produced meaningful clinical outcomes and ensured the safety of participants. Before the first rTMS session, resting motor threshold (RMT) will be measured through single-pulse TMS for all participants. RMT is measured as the minimum amount of stimulation required to observe a visible movement of the abductor pollicis brevis (APB) muscle.

For the HFL rTMS group, the area of stimulation will be five cm anterior to the APB site. The intensity of stimulation will be 100% of the RMT, and there will be 30 stimulation trains per session with 5 s duration and an inter-train interval of 25 s. The localization process for the LFR rTMS group was not mentioned in the Fitzgerald study. Thus, the area of stimulation in this study will reference another randomized controlled trial by Rao et al. (2019) studying the antidepressant effect of LFR rTMS, in which the team stimulated the F4 region based on the International 10–20 System for Electrode Placement. This study yielded optimal clinical outcomes with 87% of response rate and 67% of remission rate using the same criteria as this study (a 50% reduction in HAMD score for treatment response, and a HAMD score lower than 7 for remission). The intensity

of LFR rTMS will be 110% of the RMT. 4 trains will be given per session with duration of 180 s and inter-train interval of 30 s.

For the sham group, the study will adopt the “eSham” system developed by Borckardt et al. (2008). The “eSham” system is an improved version of sham TMS treatment that can successfully achieve scalp sensations similar to those of real rTMS, yet not influence the targeted region of stimulation. Participants in Borckardt’s study reported similar ratings on feelings of pain, tingling, piercing, pinching, etc. for both real rTMS and sham rTMS treatments. The research conducted by Rao et al. (2019) used this system and has proven its blinding effect. When asked which group patients believed they belonged to, the majority of subjects (26/30 patients) thought they were receiving the real treatment after the first session. Most of them did not change their beliefs when asked the same question after the last session (the 20th treatment), in which 20/30 patients believed they received the active treatment. Therefore, the “eSham” system should be effective in subject blinding.

The patients will have treatment sessions 5 times a week during weekdays for 4 weeks, thus 20 sessions in total. The severity of patients’ symptoms will be assessed by clinicians based on HAMD at the first day before their first rTMS treatment, and the scores will be recorded as baseline data (Baseline assessment of MDD). HAMD scores will be measured again right after the last session in the fourth week (Endpoint assessment of MDD). A follow-up data report of HAMD scores will be collected 6 months after the end point measurement (6-month follow-up assessment of MDD) to assess the long-term antidepressant effect.

Data Analysis

Measures

- **Baseline assessment of MDD:** initial HAMD scores recorded before the first rTMS session.
- **Endpoint assessment of MDD:** HAMD scores recorded after the last rTMS session.
- **6-month follow-up assessment of MDD:** HAMD scores recorded 6 months after the last rTMS session.
- **Response rate:** % of patients with (endpoint HAMD - baseline HAMD)/baseline HAMD > 50%.
- **Remission rate:** % of patients with endpoint HAMD < 7.
- **Long-term symptom reduction rate:** % of patients with (6-month follow-up HAMD - baseline HAMD)/baseline HAMD > 40%.
- **Percentage of younger (age 20-50) responders:** (responders in younger age group/total younger participants in the group) *100%.
- **Percentage of older (age 50-80) responders:** (responders in older age group/total older participants in the group) *100%.

Statistical Analysis

To compare the differences in response rate, remission rate, and long-term response rate of HFL rTMS and LFR rTMS, a one-way analysis of variance (ANOVA) will be performed. A Tukey’s range test will be applied afterwards to

assess if the differences are statistically significant. To examine how age influences the outcomes after patients receive the two rTMS variations, a two-factor ANOVA will be used with the first factor being the type of treatment applied and the second factor being age. Tukey's test will be performed again to test the statistical significance, similar to the analysis after the one-way ANOVA.

Conclusion

This paper reviews previous literature comparing the efficacy of HFL rTMS and LFR rTMS in the treatment of MDD, while also examining other factors related to the clinical outcomes and limitations of past studies. It is suggested that LFR rTMS can be a potential alternative to HFL rTMS due to its lower cost, higher public accessibility, and reduced amount of discomfort and side effects. An experiment with 150 subjects is also proposed in the paper to explore the remission rates, long-term response rates, and the role of patients' age with respect to treatment outcome for HFL and LFR rTMS. I hypothesize that both active treatment groups will have better treatment results compared to the sham group. Younger patients will respond better than the elders in both rTMS treatments, but the elder group receiving LFR rTMS will respond better than the group receiving HFL rTMS. If the results demonstrate that both HFL and LFR rTMS are effective treatment methods for MDD with comparable remission rates, long-term effects, and age ranges responding to the treatment, the study provides more evidence indicating the practicality of adopting LFR rTMS as a replacement for HFL rTMS.

References

- Barlow, C., & Targum, S. D. (2007). Hippocampal neurogenesis: Can it be a Marker for new antidepressants?. *Psychiatry*, *4*(5), 18–20.
- Baxter, L. R. (1989). Reduction of prefrontal cortex glucose metabolism common to three types of depression. *Archives of General Psychiatry*, *46*(3), 243. doi:10.1001/archpsyc.1989.01810030049007
- Berlow, Y. A., Zandvakili, A., & Philip, N. S. (2020). Low frequency right-sided and high frequency left-sided repetitive transcranial magnetic stimulation for depression: The evidence of equivalence. *Brain Stimulation*, *13*(6), 1793-1795. doi:10.1016/j.brs.2020.10.005
- Borckardt, J. J., Walker, J., Branham, R. K., Rydin-Gray, S., Hunter, C., Beeson, H., Reeves, S. T., Madan, A., Sackeim, H., & George, M. S. (2008). Development and evaluation of a portable sham transcranial magnetic stimulation system. *Brain stimulation*, *1*(1), 52–59. <https://doi.org/10.1016/j.brs.2007.09.003>
- Brunoni, A. R., Chaimani, A., Moffa, A. H., Razza, L. B., Gattaz, W. F., Daskalakis, Z. J., & Carvalho, A. F. (2017). Repetitive transcranial magnetic stimulation for the acute treatment of major depressive episodes: A systematic review with network meta-analysis. *JAMA psychiatry*, *74*(2), 143–152. <https://doi.org/10.1001/jamapsychiatry.2016.3644>

- Cao, X., Deng, C., Su, X., & Guo, Y. (2018). Response and remission rates following high-frequency vs. low-frequency repetitive transcranial magnetic stimulation (rTMS) over right DLPFC for treating major depressive disorder (M): A meta-analysis of randomized, double-blind trials. *Frontiers in Psychiatry, 9*. doi:10.3389/fpsy.2018.00413
- Carpenter, L. L., Janicak, P. G., Aaronson, S. T., Boyadjis, T., Brock, D. G., Cook, I. A., . . . Demitrack, M. A. (2012). Transcranial magnetic stimulation (TMS) for major depression: A multisite, naturalistic, observational study of acute treatment outcomes in clinical practice. *Depression and Anxiety, 29*(7), 587-596. doi:10.1002/da.21969
- Caulfield, K. A., & George, M. S. (2020). Treating the mental health effects of COVID-19: The need for At-home neurotherapeutics is now. *Brain Stimulation, 13*(4), 939-940. doi:10.1016/j.brs.2020.04.005
- Chen, L., Hudaib, A., Hoy, K. E., & Fitzgerald, P. B. (2020). Efficacy, efficiency and safety of high-frequency repetitive transcranial magnetic stimulation applied more than once a day in depression: A systematic review. *Journal of Affective Disorders, 277*, 986-996. doi:10.1016/j.jad.2020.09.035
- Fitzgerald, P. B., & Daskalakis, Z. J. (2011). The effects of repetitive transcranial magnetic stimulation in the treatment of depression. *Expert Review of Medical Devices, 8*(1), 85-95. doi:10.1586/erd.10.57
- Fitzgerald, P., Fountain, S., & Daskalakis, Z. (2006). A comprehensive review of the effects of rtms on motor cortical excitability and inhibition. *Clinical Neurophysiology, 117*(12), 2584-2596. doi:10.1016/j.clinph.2006.06.712
- Fitzgerald, P. B., Hoy, K. E., & Daskalakis, Z. J. (2020). Left handedness and response to repetitive transcranial magnetic stimulation in major depressive disorder. *The World Journal of Biological Psychiatry, 22*(4), 310-314. doi:10.1080/15622975.2020.1795255
- Fitzgerald, P. B., Hoy, K., Daskalakis, Z. J., & Kulkarni, J. (2009). A randomized trial of the anti-depressant effects of low- and high-frequency transcranial magnetic stimulation in treatment-resistant depression. *Depression and Anxiety, 26*(3), 229-234. doi:10.1002/da.20454
- Janicak, P. G., Nahas, Z., Lisanby, S. H., Solvason, H. B., Sampson, S. M., McDonald, W. M., . . . Schatzberg, A. F. (2010). Durability of clinical benefit with transcranial magnetic stimulation (tms) in the treatment of pharmacoresistant major depression: Assessment of relapse during a 6-month, multisite, open-label study. *Brain Stimulation, 3*(4), 187-199. doi:10.1016/j.brs.2010.07.003
- Kaur, M., Michael, J. A., Fitzgibbon, B. M., Hoy, K. E., & Fitzgerald, P. B. (2019). Low-frequency rtms is better tolerated than high-frequency rtms in healthy people: Empirical evidence from a single session study. *Journal of Psychiatric Research, 113*, 79-82. doi:10.1016/j.jpsychires.2019.03.015
- Kayser, S., Bewernick, B., Hurlmann, R., & Schlaepfer, T. (2012). P.2.a.025 Similar seizure characteristics of electroconvulsive therapy and magnetic seizure therapy in treatment-resistant depression. *European Neuropsychopharmacology, 22*. doi:10.1016/s0924-977x(12)70354-1

- Kedzior, K. K., Azorina, V., & Reitz, S. K. (2014). More female patients and fewer stimuli per session are associated with the short-term antidepressant properties of repetitive transcranial magnetic stimulation (rTMS): a meta-analysis of 54 sham-controlled studies published between 1997-2013. *Neuropsychiatric disease and treatment*, *10*, 727–756. <https://doi.org/10.2147/NDT.S58405>
- Kedzior, K. K., Reitz, S. K., Azorina, V., & Loo, C. (2015). Durability of the antidepressant effect of the high-frequency repetitive transcranial magnetic stimulation (rTMS) in the absence of maintenance treatment in major depression: A systematic review and meta-analysis of 16 double-blind, randomized, sham-controlled trials. *Depression and Anxiety*, *32*(3), 193-203. doi:10.1002/da.22339
- Lefaucheur J-P, Aleman A, Baeken C, Benninger DH, Brunelin J, Lazzaro VD, et al. Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS): an update (2014e2018). *Clin Neurophysiol* 2020. <https://doi.org/10.1016/j.clinph.2019.11.002>.
- Leucht, S., Fennema, H., Engel, R., Kaspers–Janssen, M., Lepping, P., & Szegedi, A. (2013). What does the HAMD mean?. *Journal of Affective Disorders*, *148*(2-3), 243-248. doi:10.1016/J.JAD.2012.12.001
- Luan, D., Zhao, M. G., Shi, Y. C., Li, L., Cao, Y. J., Feng, H. X., & Zhang, Z. J. (2020). Mechanisms of repetitive transcranial magnetic stimulation for anti-depression: Evidence from preclinical studies. *World journal of psychiatry*, *10*(10), 223–233. <https://doi.org/10.5498/wjp.v10.i10.223>
- Magpro r30. (n.d.). Retrieved September 25, 2021, from <https://www.magventure.com/component/k2/stimulators/magpro-30>
- Miron, J., Sheen, J., Mansouri, F., Blumberger, D. M., Daskalakis, Z. J., Vila-Rodriguez, F., & Downar, J. (2020a). The role of low-frequency repetitive transcranial magnetic stimulation in major depression: A call to increase the evidence base. *Brain Stimulation*, *13*(5), 1296-1297. doi:10.1016/j.brs.2020.06.013
- Miron, J., Voetterl, H., Mansouri, F., Blumberger, D. M., Daskalakis, Z. J., & Downar, J. (2020b). A case series of a novel 1 Hz right-sided dorsolateral prefrontal CORTEX rTMS protocol in major depression. *Brain Stimulation*, *13*(2), 372-374. doi:10.1016/j.brs.2019.11.006
- Mogg, A., Pluck, G., Eranti, S., Landau, S., Purvis, R., Brown, R., . . . McLoughlin, D. (2008). A randomized controlled trial with 4-month follow-up of adjunctive repetitive transcranial magnetic stimulation of the left prefrontal cortex for depression. *Psychological Medicine*, *38*(3), 323-333. doi:10.1017/S0033291707001663
- Pallanti, S., Cantisani, A., Grassi, G., Antonini, S., Cecchelli, C., Burian, J., . . . Quercioli, L. (2012). RTMS age-dependent response in treatment-resistant depressed subjects: A mini-review. *CNS Spectrums*, *17*(1), 24-30. doi:10.1017/S1092852912000417
- Peng, Z., Zhou, C., Xue, S., Bai, J., Yu, S., Li, X., Wang, H., & Tan, Q. (2018). Mechanism of repetitive transcranial magnetic stimulation for depression. *Shanghai archives of psychiatry*, *30*(2), 84–92. <https://doi.org/10.11919/j.issn.1002-0829.217047>

- Rao, V., Bechtold, K., McCann, U., Roy, D., Peters, M., Vaishnavi, S., . . . Reti, I. (2019). Low-frequency right repetitive transcranial magnetic stimulation for the treatment of depression after traumatic brain injury: A randomized sham-controlled pilot study. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 31(4), 306-318. doi:10.1176/appi.neuropsych.17110338
- Rosa, M. A., Gattaz, W. F., Pascual-Leone, A., Fregni, F., Rosa, M. O., Rumi, D. O., . . . Marcolin, M. A. (2006). Comparison of repetitive transcranial magnetic stimulation and electroconvulsive therapy in unipolar non-psychotic refractory depression: A randomized, single-blind study. *The International Journal of Neuropsychopharmacology*, 9(06), 667. doi:10.1017/s1461145706007127
- Sabesan, P., Lankappa, S., Khalifa, N., Krishnan, V., Gandhi, R., & Palaniyappan, L. (2015). Transcranial magnetic stimulation for geriatric depression: Promises and pitfalls. *World journal of psychiatry*, 5(2), 170–181. <https://doi.org/10.5498/wjp.v5.i2.170>
- Schutter, D. (2009). Antidepressant efficacy of high-frequency transcranial Magnetic stimulation over the left dorsolateral prefrontal cortex in double-blind sham-controlled designs: A meta-analysis. *Psychological Medicine*, 39(1), 65-75. doi:10.1017/S0033291708003462
- Valiengo, L., Benseñor, I., Lotufo, P., Fraguas Jr, R., & Brunoni, A. (2013). Transcranial direct current stimulation and repetitive transcranial magnetic stimulation in consultation-liaison psychiatry. *Brazilian Journal of Medical and Biological Research*, 46(10), 815-908. doi:10.1590/1414-431x20133115



#ChallengeAccepted for Women in Turkey: Trending Instagram Participation or Collective Feminist Activism?

Yong Jie (Justina) Chou

Author Background: *Yong Jie (Justina) Chou currently attends St. Paul's Co-educational College in Hong Kong SAR, China. Her Pioneer research concentration was in the field of sociology/media studies and titled "Social Media and Human Progress."*

Abstract

Through this paper, I take the #ChallengeAccepted hashtag campaign in 2020 where female celebrities and influencers shared black-and-white selfies on Instagram as a case study to examine its ramifications as a response against sexual violence in Turkey and confront the question: is a viral hashtag campaign truly effective in organizing feminist collective activism? To discuss the effects and limitations of the movement in developing a feminist counter-public in Turkey, I first reviewed existing studies on hashtag as a mobilizing tool for online feminist causes. Following this, I analyzed the content of Instagram posts participating in the #ChallengeAccepted campaign, identified their key features, and compared them with other feminist hashtag movements of a similar scale on Twitter. Through my analysis, I evaluated the role of platform politics of Instagram in causing the reappropriation of images, and suggested that the short-lived hashtag campaign could not retain its Turkish symbol of black-and-white photographs of women of death amidst hashtags, nominations and artful selfies. In considering the flaws of viral social advocacy, I hope to present solutions for visually-oriented social media platforms to promote, unify and sustain feminist causes against sexual violence to achieve social justice.

1. Introduction

Hashtag campaigns on social media have provided women with platforms to resist discrimination in digital campaigns such as #WhyIStayed and #MeToo by allowing women to voice their personal experiences of being harassed or harmed and enabling them to show support to each other. While some movements have gained widespread attention and brought renewed changes, campaigns on social media are not without their flaws. Some hashtag campaigns have been labelled as

“slacktivism” since they require low personal investment like pressing a like or a share button. Slacktivism, the practice of only supporting a political or social cause through social media or online petitions, is characterized as online activism involving little effort or commitment (Cabrera, 2017, p.1). One of the most heavily critiqued campaigns was #Bringbackourgirls in 2014. It aimed to bring back 216 schoolgirls who were abducted, yet many girls remained missing one year after the hashtag trended on social media.

Hashtags that have been previously put to use in other campaigns may strengthen their viral spread-ability and visibility in the algorithm’s sorting on social media, as a metadata tag that allows cross-referencing of content shared under the same topic or theme (Chang & Iyer, 2012, p.251). In July 2020, Instagram feeds were inundated with over 4.5 million black-and-white pictures of women (Jamal, 2020) captioned with words of female empowerment and hashtags of #ChallengeAccepted or #womensupportingwomen. The death of 27-year-old Pınar Gültekin, who was brutally strangled and murdered by her ex-boyfriend, led to female outrage and street protests in Turkey. The hashtag campaign on Instagram began when women started posting black-and-white pictures of themselves on social media as a protest against increasing femicide, the misogynistic killing of women in Turkey, as black-and-white photos of female victims often appeared in newspapers and on TV screens. Women continued to nominate other women and famous celebrities including Ivanka Trump, Khloe Kardashian and Paris Hilton who declared their support for the movement.

I take the #ChallengedAccepted movement on Instagram as a case study to shed light on the effects of recent viral hashtag campaigns as a response to femicide. The purpose of my research is to answer the following questions: to what extent is an Instagram viral hashtag trend effective in organizing feminist collective activism against sexual violence in Turkey? How do hashtag challenge campaigns on Instagram compare to other text-based social media as a mobilizing instrument for the development of feminist counter-publics? Few studies after 2015 have analyzed the increasing salience of visually-oriented social media, namely Instagram, as a resource for feminist activism among young people. Furthermore, as an active user on Instagram, I was personally moved by the black-and-white images and prompted to take action by sharing the posts on my own account. Hence, I am genuinely interested in investigating different aspects of the movement and factors which bring success to movements and their possible limitations.

In my paper, I will first review existing studies on Instagram hashtag campaigns as a mobilizing tool for online actions and their implications on combatting sexual violence in Turkey. I will analyze the content (including the photo and caption) of #ChallengeAccepted posts created by celebrities and discuss their effects and limitations in bringing visibility to a pre-existing feminist counter-public developed by website, blogs and feminist organizations based in Turkey, in which shared identities and common narratives are built on platforms comprising elements of the public sphere. I will also compare the hashtag movement to another digital feminist movement of similar scale, #NiUnaMenos, to show how hashtag campaigns play out on Instagram as opposed to text-based platforms in terms of spreading important messages. I will argue that due to the limitations of hashtag politics and the pitfalls of platform politics on Instagram, the #ChallengeAccepted campaign effectively created concern for femicide, but it

largely relied on influences on social media directing attention back to femicide in Turkey. Nevertheless, the hashtag campaign enhanced visibility of an existing feminist counter-public sphere on social media.

My research intends to rethink the affordances of viral hashtag campaigns on Instagram and other emerging visually-oriented social media platforms to promote, unify and sustain feminist causes against gender-based violence in Turkey. I hope to contribute to an understanding of how hashtag activism on increasingly popular visually-oriented platforms on socio-political issues among young people can be used to develop a feminist counter-public, and create a basis for further research on discursive techniques employed for cyber-activists on advancing feminist movements against sexual violence.

2. Literature Review

2.1 Contextualizing Activism against Gender-based Violence and Femicide in Turkey

Gender-based violence refers to physical, verbal, or emotional harmful acts directed at an individual based on their gender (UNHCR, n.d.). It has been a prominent issue worldwide, especially in Africa and the Eastern Mediterranean Region. Throughout Turkey, women are subjected to violence as a result of several factors including lack of proper law enforcement, lack of government support for gender equality, and patriarchal customs (Zhang, 2015, p.432). The practice of honor killing, the murder of a woman or girl by male family members, prevails in the patriarchal Turkish society, where the killers' actions are justified by claiming that women have brought dishonor upon family prestige by engaging in "sexually immoral actions" (Dailey, n.d.). Today, women in Turkey still live under the threat of gender-based violence and femicide despite legal regulations made to prevent violence against women. Moreover, femicides are underreported and women tend to stay silent so as to conform to the male-dominated culture in Turkey. The number of women killed has been on the rise, with 474 women murdered in 2019, though the Turkish government has not worked to keep records of violence against women (Thelwell, 2020).

Turkish feminists have been actively raising awareness on social media since the early 2010s, through disseminating information of protests, and gathering international support through hashtag campaigns like #ThereisamassacreofWomen and #istanbulsözleşmesiyadır (*say no to violence against women*). Websites and blogs have been established for women to connect with each other and share their personal experiences. Online campaigns such as *My Body My Choice* were organized towards the initiative to ban abortion in 2012, reaching an international audience and gaining support. Feminist organizations like the We Will Stop Femicide platform have provided legal assistance to women in danger and have organized campaigns to support laws protecting women's rights. The Istanbul Convention, an international human rights treaty that specifically targets gender-based violence, used to enhance law enforcement and protect women from femicide in the country, until the Turkish government's withdrawal from it in 2021.

2.2 Instagram as a Digital Feminist Activist Platform

The reliance of Instagram on visual content attracts more feminists to produce images and videos to disseminate messages to the younger population and mobilize actions. Scholars have argued that visual images of street protests on social media may have an effect of provoking emotions and advancing movements, such as ones in the 2011 Egyptian Revolution (Ogan & Bas, 2020, p.5559). The capability of Instagram as a photo and video sharing platform allows “bright digital protest flyers, stylized illustrated portraits” to be published as opposed to “dull block texts” on Twitter or blogs (Stewart & Ghaffary, 2020), making it a natural strong fit for creating awareness for various social justice issues.

However, one danger of Instagram in promoting social causes is the reappropriation of images into growing aesthetic content. Some artists have specifically chosen Instagram as their means of expression regarding gender and feminist issues, such as fashion model Adwoa Aboah and actress Jameela Jamil who launched their respective Instagram pages *Gurls Talk* and *iweigh* to spread feminist messages. Through establishing and assigning aesthetic values to feminist causes, feminist dialogues may transform into marketing material for celebrities that they can use to appeal to their target audience with similar values and beliefs. This leads to a perceived difficulty between balancing creative aesthetics and maintaining radical engagement with audiences (Crepax, 2020, p.72) through a call-to-action.

The major disparity between hashtag campaigns on Instagram and other popular text-based platforms lies in Instagram’s platform architectures as a business and advertising model. As a social media with algorithms favoring commercial advertisements or sponsored content (Sener, 2021), Instagram is easily commodified into an advertising platform, with unlimited hashtags being used to promote content, increase exposure and audience reach. Emily Stewart and Shirin Ghaffary (2020) have described it as “the concept of having race, class, gender, and other identity markers overlap and factor into the usual funny memes, skin care routines, and fitness videos”. Taking the example of the Body Positive Movement, Cwynar-Horta’s research (2016) reached the conclusion that feminist activism and movements on Instagram involving female social media influencers can lead to a situation where corporations begin to commodify fashion bloggers who present themselves as “body advocates.” The sampled data concerning the use of hashtags and images of Instagram posts are associated with content promotion of yoga classes, sponsorship of sportswear and much more (Cwynar-Horta, 2016). Twitter users, on the other hand, are advised by platform guidelines to not use more than two hashtags in order to ensure communication and allow activists to express themselves authentically without obstructing the dissemination of messages and discussions.

Although the new function of sharing a longer video as an IGTV was launched on Instagram in 2018 as a tool for sharing more complex messages and initiating more thorough conversations between users, as opposed to text-based sites like Twitter and Reddit, where polarized debates frequently occur, Instagram is not designed to facilitate socio-political conversation. On Twitter, activists and dissidents can retweet by typing their own comments under the threads as a way to spread their message; on Reddit, subreddits can be formed quickly without the need for a new identity, and many conversations can be self-

regulated without assistance. Both rely on unambiguous text messages for communication, as opposed to highly personalized images and videos, which may be repurposed or perceived in a different way, used to engage audiences on Instagram.

2.3 Implications of the #ChallengeAccepted

Studies investigating challenge-based hashtag campaigns on social media with purpose of raising awareness for social justice projects in the past have proposed several features of contagious content. For instance, the Ice Bucket Challenge, which encouraged nominated people to pour a bucket of ice water over their heads, went viral on social media during the summer of 2014 and successfully gathered a donation of more than US \$220 million for medical research in the amyotrophic lateral sclerosis disease (time.com., 2020). According to Mahoney's research (2015), the motivating factors for participation in challenge-based hashtag campaigns include direct nomination and social pressure, as nominations mainly took place within the highly visible social sphere on Facebook pages. A nominated non-participant of the Ice Bucket Challenge had stated: "I also didn't like the peer pressure side of it. It also became a middle school 'who is my best friend?'" (Mahoney, 2015, p.75). Social currency of the nominated to look smart and socially desirable in front of their friends and followers has been attributed as the factor behind the success of viral content spread on social media (Pressgrove & McKeever & Jang, 2018, p.2), instead of practical values, personal devotion, and emotional attachment to the social justice issue.

Despite the spread-ability of the hashtag challenges through interpersonal channels of personal networks on social media (Mogensen, 2014), there are negative outcomes of the call-out strategy. Firstly, the hashtag challenges' long-term viability is questionable: the Ice Bucket Challenge on Facebook died off after the summer, while participation in the #ChallengeAccepted campaign dropped dramatically in a matter of weeks. Secondly, hashtags like #ChallengeAccepted, which do not include explicit references to a specific social context, are readily hijacked and manipulated. The initial use of the hashtag #ChallengeAccepted can be dated back in 2016 when it was used to raise cancer awareness and spread positivity. Instagram users posted black-and-white pictures of themselves and nominated others through a similar call-out strategy. It is unclear as to where and how the hashtag challenge was launched again in the summer of 2020 as a campaign for female empowerment (Jamal, 2020). Most sources indicate that the hashtag challenge campaign was set up to bring attention to the high rates of femicide in Turkey, while some suggest that the cycle of the challenge was initiated by Brazilian journalist Ana Paula Padrão who posted a monochromatic photograph of herself with the hashtag #womensupportingwomen (Moss, 2020).

Given Instagram's platform features and the nature of challenge-based hashtag campaigns, short-lived hashtag movements on Instagram are expected to accomplish impressive results in bringing a social issue to the surface while being unable to influence large-scale social developments.

3. Methodology

This study aims to analyze the #ChallengeAccepted campaign on Instagram in regard to developing a feminist counter-public to combat femicide in the Turkish context. The methodology of this study is a primary content analysis of a handful of #ChallengeAccepted posts made by renowned celebrities, as well as secondary analysis of quantitative data from existing papers. I will also look at the efforts seeking to bring back discussion to the femicide situation in Turkey and evaluate their visibility.

3.1 Content Analysis of Celebrity Posts

Taking the form of a virtual protest, women could participate in the “challenge” by posting a black-and-white selfie or photo of themselves with the hashtags #ChallengeAccepted, #womensupportingwomen and Turkish hashtags related to the Istanbul Convention like #istanbulsözleşmesiyaşatır. Ustel (2021) noted in her article that participants were also encouraged to send the following prompt to their female friends and family through the direct message function on Instagram:

I was careful to choose who I think will meet the challenge, but above all who [sic] I know who shares this type of thinking, among women there are several criticisms; instead, we should take care of each other. We are beautiful the way we are. Post a photo in black and white alone, write “challenge accepted” and mention my name. Identify 50 women to do the same, in private. I chose you because you are beautiful, strong and incredible. Let’s [heart emoji] each other! (Ustel, 2021, p.38)



Figure 1. @khloekardashian's #ChallengeAccepted Instagram post (Kardashian, 2020)



Figure 2. @jennifersaniston's #ChallengeAccepted Instagram post (Aniston, 2020)

In the monochromatic selfie of Khloe Kardashian [*@khloekardashian*], an American socialite and model, she makes the statement, “To all my Queens - Let’s spread love” with the hashtag *#womensupportingwomen*. The post accumulated almost 2 million likes and just under 10 thousand comments until August 2021, reinforcing the use of social media in building social empowerment. American actress and TV producer Jennifer Aniston [*@jenniferaniston*] posted a black-and-white portrait of herself captioned with the hashtag *#challengeaccepted*. She also encourages her audience to “REGISTER TO VOTE for the issues that affect women.” The capitalized message of support has attracted more than 4 million likes and more than 20 thousand comments under the post.



Figure 3. *@nicolaanpeltz's #ChallengeAccepted Instagram post (Peltz, 2020)*



Figure 4. *@serenayyss's #ChallengeAccepted Instagram post (Sarıkaya, 2020)*

American actress Nicola Peltz [*@nicolaanpeltz*] applied a black-and-white filter to the photograph of herself, posing with a pair of sunglasses in a crop top on the beach. In her caption, she nominated Victoria Beckham (her mother-in-law) to participate in the challenge and expressed gratitude to “have so many strong women in [her] life that empower[s] [her] and those around them.” On the other side of the globe, Turkish actress and model Serenay Sarıkaya [*@serenayyss*] captioned her black-and-white selfie with three hashtags: *#challengeaccepted*, *#womensupportingwomen* and *#istanbulsözleşmesiyaşatır*,

transmitting a simple yet powerful message of motivation and support to promote solidarity with Turkish women.

3.2 Secondary Analysis

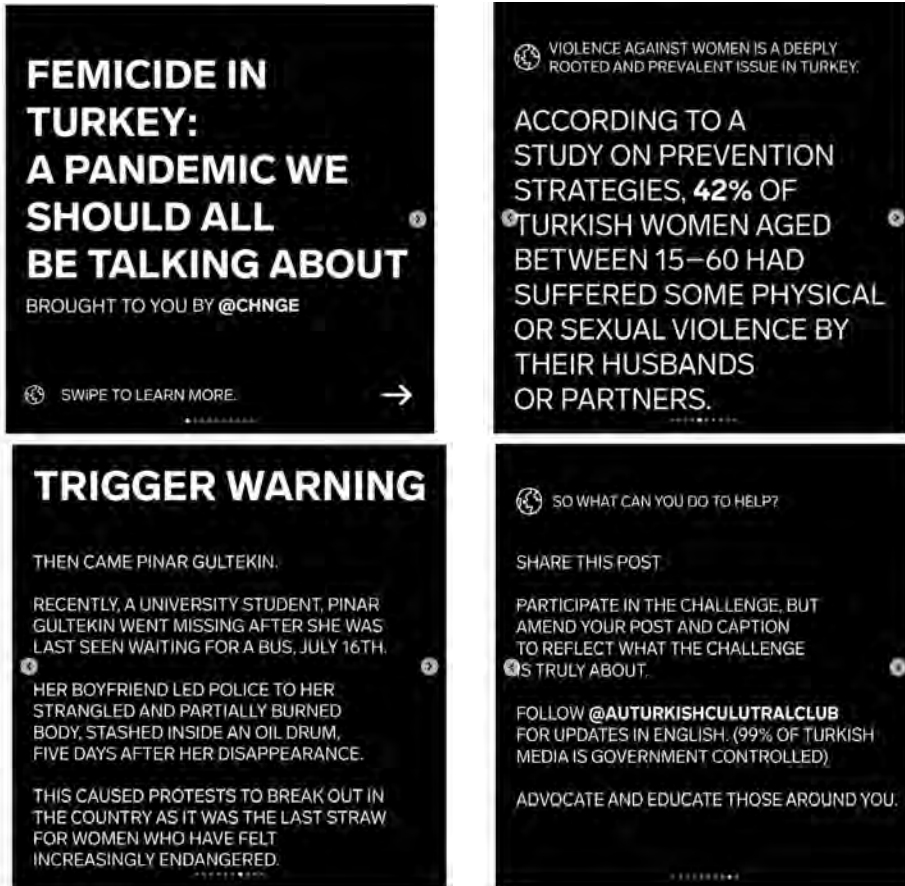
ÖZŞENLER's paper (2021) selected 51 posts by local and foreign women through a random sampling and grouped them under themes of messages with aims of support, courage, motivation, power, energy and leadership status. Out of the 51 posts, 26 #womensupportingwomen hashtags were used after the most-used hashtag of #ChallengeAccepted, instead of those related to the Istanbul Convention or femicide in Turkey. The most frequent aims of posts are support, courage, motivation, and the importance of embracing diversity.

Yet, it is argued that the campaigns are mostly limited to personal action apart from being indicative of the broader issue of femicide in Turkey (Eler, 2020). It is interesting to note that most celebrities' posts, three out of the four above, do not contain any references to femicide in Turkey, nor use the Turkish hashtag #istanbulsözleşmesiyaşatır to create concern for the femicide issue in Turkey. Jennifer Aniston has even admitted that she did not know where and how this challenge came from, stating that "but who doesn't love a good reason to support women!" (Aniston, 2020) This shows that most #ChallengeAccepted posts created by western celebrities were in fact limited to messages of female empowerment and did not allow audiences to learn about rising femicide rates in Turkey.

These photographs of women have also been described as "posed, filtered, often taken from flattering angles and accompanied by benign captions" (Lorenz, 2020). As an awareness campaign, the black-and-white photos of women are exceptionally beautiful, some even constituted as "hot selfie," clearly conveying the message that women are beautiful and strong. Despite the strong sense of exalted empowerment and a strong sense of sisterhood, there are few signs that the photos are to mirror black-and-white photos of murdered women splashed across newspapers and televisions in Turkey.

3.3 Influences directing attention back to Femicide in Turkey

The loss of significant messages in the campaign in raising awareness for violence against Turkish women has prompted efforts on Instagram aiming to bring attention back to the femicide issue in Turkey with criticism mainly coming from Twitter users.



Figures 5-8. Instagram post from @chnge clarifying the origin of the #ChallengeAccepted hashtag movement (CHNGE™, 2020)

For instance, CHNGE™ [@chnge], an ethical clothing brand which works with manufacturing workers in Turkey, shared a post explaining the prevalent issue of Turkish femicide and the meaning behind the black-and-white photograph hashtag challenge campaign. Through the use of statistics and a brief summary of the murder of Pinar Gültekin, the post informed the audience about the origin of the hashtag challenge, which triggered a protest both on the streets in Turkey and on social media. The post successfully reached millions of people with more than a hundred thousand likes and over 1350 comments. Audiences were also encouraged to share the information with those around them through a call-to-action.

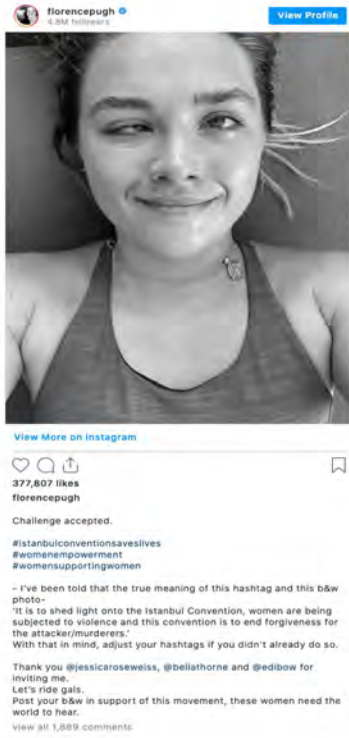


Figure 9. *@florencepugh's #ChallengeAccepted Instagram post (Pugh, 2020)*



Figure 10. *@dilanbozyel's #ChallengeAccepted Instagram post (Bozyel, 2020)*

Furthermore, some artists and influencers have made use of both the #ChallengeAccepted hashtag and ones related to the Istanbul Convention. Contrary to the majority of beautiful gestures, English actress Florence Pugh [*@florencepugh*] posted a humorous selfie of herself making a face and wrote "I've been told that the true meaning of this hashtag and this b&w photo- 'It is to shed light onto the Istanbul Convention, women are being subjected to violence and this convention is to end forgiveness for the attacker/murderers.' With that in mind, adjust your hashtags." The post using the hashtag #istanbulconventionsaveslives motivated people to spread the word for Turkish women in support of the Istanbul Convention by gaining over 377 thousand likes. Istanbul-based photojournalist Dylan Bozyel [*@dilanbozyel*] posted with the hashtag #istanbulsözleşmesiyagatır sitting front and center, with #challengeaccepted further down (Eler, 2020). By commemorating the Istanbul Convention in her caption, she disseminated the convention as a symbol for female solidarity in Turkey, despite the fact that she had a smaller size of audience on Instagram compared to more famous western celebrities such as Khloe Kardashian and Ivanka Trump.

3.4 Criticism from Twitter users



Figure 11. *Twitter post by bipolar baddie [@headvches] criticizing the loss of message on femicide in Turkey in the #ChallengeAccepted hashtag movement (O'Malley, 2020)*

Some Twitter users have also voiced their criticism in regard to the limitations of “challenge accepted selfie” images in conveying the message of enlisting support for the femicide issue in Turkey. It may be inferred that they believe text-based platforms are more effective in spreading information about the Turkish origin of the hashtag movement to users who are still unaware of it.

4. Analysis and Discussion

4.1 Reappropriation of #ChallengeAccepted Selfies in Celebrity Posts

A number of renowned celebrities supported the #ChallengeAccepted hashtag campaign. The movement became viral on Instagram when women started nominating one another to partake in the challenge, which in fact requires minimal effort: posting a black-and-white selfie with the hashtag #ChallengeAccepted. The bandwagon effect as a means to promote viral dissemination of feminist messages accumulated an impressive number of posts shared under the hashtag #ChallengeAccepted.

The participation of celebrities in the movement contributes to performance activism, as well as the commodification of aesthetics and celebrity publicity. For example, American filmmaker Ava DuVernay [@ava], who has 2.2 million followers on Instagram, shared a photograph of herself captioned “pretty sure this is just a cool excuse to post pix” (DuVernay, 2020). The post has gained 50 thousand likes, drawing criticism because some believe the idea of the hashtag challenge as a way to change patriarchal beliefs or practices against gender equality is not persuasive. According to the idea of celebrity attention philanthropy, celebrities were perhaps more concerned with constructing individual narratives (Nabil, p.5) and philanthropic social capital of supporting social causes without investing in higher cost actions such as participating in

petitions and initiating donations to address sexual violence against women in Turkey.

The artful selfies in the hashtag campaign may have different meanings in different contexts. In contrast to some western celebrities who post their photographs to promote art and fashion, women in Turkey believed the most important reason for participating in the challenge was the black-and-white pictures of women as “a reference to photos that mourners pin to their clothes at the funeral of a murdered woman” (Eler, 2020). As seen in the celebrity posts, the message of support in the Turkish context and the symbol of black-and-white photographs of women of death were mostly buried in the surge of hashtags, nominations, and artful selfies.

Moreover, the hashtag campaign’s role in organizing a collective feminist cause relied heavily on efforts directing attention back to the issue of femicide in Turkey by social media users as opposed to influential celebrities. Most Turkish women and artists used hashtags such as #IstanbulSozlesmesiYasatir (*The Istanbul Convention helps women stay alive*) and #istanbulsözleşmesiyaşatır to protest against potential government plans of withdrawal from the agreement, while only a small number of US celebrities shared the Turkish-language hashtag about the Istanbul Convention and expressed support for the Turkish campaign amongst the 5 million photos shared on Instagram bearing the hashtag #ChallengeAccepted (By News from Elsewhere, 2020). Western celebrities, such as Khloe Kardashian and Kristen Bell, who have obtained a greater number of likes and follows and have garnered wider global attention than local celebrities in Turkey, did not use any hashtags related to the femicide crisis in Turkey other than #ChallengeAccepted. As the medium of spreading messages, as opposed to plain texts and call-to-actions, images are often ambiguous and subject to interpretations. These elements of an image-based hashtag campaign on Instagram contributed to the loss of significance of the black-and-white photographs of women as a form of protest when the viral trend reached audiences in the west who posted attractive selfies accompanied by messages of female empowerment. As per criticism presented by Twitter users, the hashtag campaign may have been more effective in addressing femicide in Turkey if words instead of images had been used to promote awareness of the issue on Twitter or other text-based social media.

4.2 Examining the Impact of #ChallengeAccepted on Femicide in Turkey

The #ChallengeAccepted movement raised some awareness for the femicide issue in Turkey but was less effective in supporting female victims of gender-based violence in tangible ways compared to the content by feminist organizations on other social media.

Out of the three major types of Turkish feminist organizations identified by Şener (namely informal and primarily web-based small initiatives on blogs or websites, large platforms and networks mobilizing both online and offline, and formal associations) (Şener, 2021, p.18), one of the most well-known is We Will Stop Femicide Platform. In addition to its website, it is active on social media sites including Instagram, Facebook and Twitter, and mainly serves women in Turkey by providing a hotline for female victims of gender-based violence. In addition to

providing legal assistance and joining court cases to support struggling victims, it leads feminist movements in Turkey by collecting the data of femicide cases and publishing those statistics on a monthly basis, which the Turkish government has refused to do. It also organizes publicity meetings at schools, universities and establishes branches to expand the number of supporters and coordinate street protests when necessary (*We Will Stop Femicide Platform, n.d.*).

Compared with hashtag campaigns on other social media in Turkish digital feminist activism, Instagram is used as a platform to publicize activities and events of feminist organizations through affective photos and videos, and engage with younger audiences through visual materials such as GIF, pictures and graphics (Şener, 2021, p.17), raising awareness. On the other hand, there is an alternative public space on Twitter in which the functions of retweet and mentions allow feminist messages from hashtag campaigns to spread rapidly and extensively, thus contributing to the accelerating mobilization of protests (Şener, 2021, p.15). Moreover, feminist groups on Twitter provide hotline services for emergency contact. In some cases, followers of Twitter pages have sought for help. The Twitter groups directed these women to professional associates, successfully using social connections on social media to produce concrete offline actions to help Turkish women in-need (Şener, 2021, p.19). Twitter is also known as a digital platform where women create an agenda to prevent the Turkish authorities' from further restricting women's rights through their critical discussion of controversial policies and personal experiences of sexual violence in real life.

The #ChallengeAccepted hashtag movement on Instagram has its place in relation to the current Turkish context as an unexpected successful awareness campaign that has aroused widespread global attention to rising femicide numbers in Turkey as a social justice issue. However, some local feminists doubted its effects in fighting against what is happening in Turkey. Maria Cerce, the coordinator of Women's Resource Centre expressed in an interview

“Personally, I think that the trend is a form of slacktivism because it only showed that people were involved,” said Celene Yang, a fourth-year biology student. “I believe that if you're going to post one of these, it will be best to have a longer caption than just 'challenge accepted' and a hashtag” (Smith, 2020).

The hashtag challenge movement may even have increased women's vulnerability and created a backlash effect for the Turkish government to announce its official decision of withdrawing from the Istanbul Convention in the summer of 2021. According to media reports, the retention of critical positions in the government by ultra-conservative rights in the recent decade has allowed eclipses by ideas of cultural distinctiveness and traditions (Özlem, 2021). In response to pro-women movements, the ruling Justice and Development Party (AKP) has increased references to the safeguarding of Turkish customs, by which they claimed to be “defending family values.” As a result, while the #ChallengeAccepted movement on Instagram has effectively drawn public attention towards violence against women in Turkey, the government's declaration may be considered as a reaction to top-down initiatives connected to globalization, as the time of announcement of a rejection of current regulations on the Istanbul Convention was close to the hashtag campaign on Instagram.

In short, the #ChallengeAccepted hashtag campaign was undoubtedly effective in spreading messages of female empowerment around the world but may not be the most effective in terms of addressing femicide in Turkey, providing support for the victims of gender violence, and making voices heard for the Turkish government.

4.3 Effects of the Hashtag Campaign in Enhancing Visibility of a Feminist Counter-public

Rapidly gaining international attention through shared posts of famous celebrities, the hashtag campaign spread messages of female empowerment, although it was taken out of the context of femicide in Turkey and the Istanbul Convention (Ustel, 2021, p.41). Through the use of hashtags like #womensupportingwomen, the hashtag campaign brings a pre-existing feminist counter-public sphere into light.

The public sphere, according to Habermas, is a realm of social activity where public opinion may be created without the use of coercion (Habermas, 1991, p.398). People have the right to free expression, yet the public sphere on social media is commonly prone to harassment and intimidation. A counter-public is a subset of a public sphere that forms around subordinated identities, commonly defined as public spheres that bring together the excluded through mutual recognition. Social media spaces provide us an environment for a better understanding of how individual elements interact with each other (Bruns, 2015), of how “density of communications is internally organized, how it maintains boundaries and relatively larger internal cohesion in relation to the larger public” (Calhoun, 1992, p.38). According to Jackson, S.J. & Banaszczyk, S., its purposes are “to build common narratives from shared identities and experiences, challenge the reproduction of hierarchies within social movement spaces, and transform and influence broader public narratives on issues central to women’s lives” (Jackson & Banaszczyk, 2016, p.393).

Pro-feminist websites and initiatives in Turkey have been working to construct a feminist counter-public by building digital sisterhood with deeper emotional ties than the surface-level natured hashtag campaign on Instagram. Examples include *5Harfliler*, *Catlak Zemin*, *Recel-blog*, which all provide an online sphere for anonymous women to connect with each other and share their personal stories of violence and oppression experienced in the patriarchal society. Technologies have created a platform for linking private and collective memories by providing a counter-public for constructive discussion of women’s shared everyday problems rather than feminist politics (Goker, 2019, p.314), although women’s movements in the past, present and future are often commemorated. Yet these websites face the challenge of keeping the agenda present in people’s minds by means such as publishing essays and initiating discussion in the face of political suppression of feminism and gender equality. Therefore, they still acknowledged the viral spread of awareness created by the #ChallengeAccepted hashtag campaign. “Because we’re at a very desperate point, we’ll embrace everything that we think can help us,” a Turkish writer and film writer, Zeynep Dilara, stated. “ChallengeAccepted is one of those things.” (Eler, 2020).

In the #ChallengeAccepted Turkish context, the visibility of pre-existing counter-public spaces is greatly enhanced through the hashtags used by celebrities. A sense of community is reinforced through the frequent use of collective pronouns of “we” and “they” in their posts, exemplified by the four examples of celebrity posts in the methodology section. The hashtag movement has forged communication at local, national, and international levels by encouraging women to tag one another and participate in the challenge. It has connected women through their personal networks despite geographical barriers (Şener, 2021, p.3) such that a counter-public for women can be exercised.

Moreover, a counter-public is also established through the influences which directed attention back to femicide in Turkey. An example is the comment section of the post by @stopfemicides, which is analyzed in Akdemir’s Bachelor’s Thesis (2021). Through sharing their personal experiences of violence against women, users stand in support of victims in Turkey and other countries and subsequently raised awareness towards the worldwide prevalence of gender-based violence. By providing women a platform to tell their stories and be heard, the hashtag movement extended its positive impact outside of Turkey globally (Akdemir, 2021).

Lastly, the use of the #istanbulsözleşmesiyadır hashtag has become a symbol in the feminist counter-public sphere in relation to the Istanbul Convention in some European countries. The increasingly conservative Turkish government and a growing number of Eastern European countries led by right-wing leaders are attacking the Istanbul Convention, which reduces violence against women, in an attempt to preserve traditional “family values” (Brechenmacher, n.d.). The people’s indignation at the inability of the government to protect women from gender-based violence is expressed in Turkish women’s usage of #istanbulsözleşmesiyadır in the #ChallengeAccepted campaign, which links women with the shared identity and symbolized defiance against the government’s stance. The Istanbul convention’s hashtags have also been translated into several languages on Instagram including German (Renk., 2021) Arabic (Gómez, 2021), Spanish (Gómez, 2021) and Catalan (Dones Juristes, 2021) by European feminism activists and organizations on Instagram, bringing the symbol of resistance to a wider audience in countries such as Hungary, Bulgaria, and the Czech Republic, whose governments are attempting to withdraw from the human rights pact.

In short, although the #ChallengeAccepted movement brings visibility to a feminist empowerment in a counter-public sphere, it is less significant than pre-existing feminist websites which engage women in enhanced emotional closeness and community building. Due to its short-lived nature, the hashtag campaign could neither challenge social hierarchies that create everyday problems for women nor influence narratives within Turkey in accordance with Jackson’s definition (2016). Rather, it could only spread messages of female empowerment.

4.4 Sustainability of the Hashtag Movement

The #ChallengeAccepted hashtag campaign on Instagram went viral by the last week of July in 2020 and petered out in the first week of August. The short-term campaign could be considered successful in its spread-ability, yet the following

reasons may contribute to its short-term nature and inability to sustain and create impact:

Firstly, the movement itself was disorganized and scattered. Some of the celebrities who participated in the challenge were unaware of the campaign's intention, and made call-to-actions that did not address the femicide issue in Turkey. One example is Jennifer Aniston's call-to-action (Aniston, 2020) for women to register to vote for issues that affect women, which does little to help stop the Turkish government from withdrawing from the Istanbul Convention.

Secondly, the political context where the campaign was started may limit widespread usage of hashtags addressing the Istanbul Convention. The repressive Turkish government may play a role in influencing the stances of hashtag posts on the femicide issue unlike other global movements which openly criticized offenders and shifted responsibility to the government. Studies have found that there is little criticism of public officials for lack of actions amongst the #ChallengeAccepted posts and online content on websites (Ogan & Bas, 2020, p.5566), while most indicated elements of support, emotion, opinion. Possible social media surveillance by the government (Sebastian, 2019, p.42) creates fear of getting on the wrong side of the political sphere, contributing to limited efforts in organizing protests through call-to-action or initiating any other activities.

The #Niunamenos hashtag movement in Latin America is a similar anti-femicide movement which can be compared with that in Turkey. The campaign started as a slogan for a protest at the Palace of the Argentine National Congress in June 2015 against the murder of 14-year-old pregnant Chiara Paez by her boyfriend, with underlying causes of poor implementation of laws, corruption, and the government's lack of political will to achieve gender equality. What began as a hashtag on Twitter and Facebook soon evolved into a social media campaign, later spread internationally to cause mass street demonstrations in Chile, Peru, Brazil, Mexico and Spain (Goñi, 2016). The hashtag movement mainly relied on Twitter and Facebook instead of Instagram as it occurred before the surging popularity of Instagram among Gen Z after 2015. The movement achieved remarkable accomplishments, including the 2020 legalization of abortion in Argentina and the Argentine government's establishment of the requirement to collect and publish national femicide statistics, and their creation a ministry for women to address gender-based violence with better transparency.

Research on digital feminist activism in Turkey against femicide has found that the political environment for campaigns plays a significant role in affecting the effectiveness of social media as a tool for social change (Ogan, 2020, p.5569). Considering the repressive political context and the nature of hashtag campaigns on Instagram, the #ChallengeAccepted movement created concern for femicide in Turkey but had limited ability in sustaining feminist impact.

5. Conclusion

Through considering #ChallengeAccepted Instagram posts shared by western celebrities and feminist organizations, individual opinions expressed on Twitter, as well as websites and platforms established to support Turkish women, I argue that the business algorithm of Instagram, attention philanthropy of celebrities on social media, and affordances of a minimal challenge-based campaign may lead




to the danger of a reappropriation of an image-based hashtag campaign. The potential loss of a significant message as the hashtag challenge spreads across the world through social networks makes Instagram more effective in promoting viral messages than developing strong digital social bonds among advocates to prompt actual actions for addressing social issues.

Therefore, viral hashtag challenge campaigns on Instagram only serve to create concern for social justice issues by attracting attention and participation. They complement a pre-existing feminist counter-public where women share their personal stories and build digital sisterhood of strong emotional bonds but are mostly insufficient to organize collective feminist activism on their own.

My research has reached the conclusion that the degree of success of a digital hashtag movement within the country of origin is highly dependent on the affordances of hashtag campaigns on the features of a particular social media, as the outcomes of hashtag campaigns on developing a feminist counter-public sphere vary across different platforms. Due to the varied platform politics of Instagram as a visually-oriented platform in contrast to Twitter and other text-based social media, the #ChallengeAccepted campaign can be considered as an exceptional case when compared to other highly successful hashtag feminist movements in the 2010s like #Niunamenos. Instagram is argued to be an effective tool for spreading messages and raising awareness with the help of aesthetic visual content, but has limited potential for developing ideas and the mobilization of offline social causes. Without clear allusions to regional social issues, challenge-based hashtags like #ChallengeAccepted are readily manipulated and converted into messages that celebrities are encouraged to support. Other text-based social media such as Twitter, Reddit, webpages, and blogs may be more effective in forming a counter-public in which users may exchange opinions and question and rebuild global narratives around women's common issues.

This study can be further extended in its data regarding feminist advocates' reactions and views in Turkey on whether the hashtag campaign is helpful in bringing attention to femicide or serves as a distraction to existing organizational works. Further research may look into the question of sustaining hashtag movements and mobilizing actions in Turkey. In the case of #ChallengeAccepted, it is worth looking into the long-term impact of the hashtag challenge on mobilization of feminist movements in Turkey or other related social causes globally against gender-based violence. This is in hope of turning the power of the global spread of positivity on social media and the international awareness raised into concrete efforts reducing sexual violence against women. Moreover, the question of how hashtag campaigns on Instagram can be improved and whether the attention created will always rely on influences redirecting focus to the social justice issue is ultimately crucial as visually-oriented social media like Instagram may become the mainstream preference of sources of news and information among youngsters in future decades.

Works Cited

- Abrams, M. (2020, July 29). #ChallengeAccepted: Is the social media trend vapid or empowering? Retrieved September 10, 2021, from <https://www.standard.co.uk/insider/living/challenge-accepted-meaning-how-started-a4510731.html>
- Akdemir, D. F. (2021). #ChallengeAccepted–Hashtag activism against femicide in Turkey: A Content Analysis of the comments made under a post by the Instagram page 'stopfemicides' (Bachelor's thesis).
- Aniston, J. [@jenniferaniston]. (2020, July 28). Thank you to all the brilliant and beautiful women in my life who sent the most uplifting messages today. [Instagram photo]. Retrieved from <https://www.instagram.com/p/CDKzW-HDogh/>
- Bozyel, D. [@dilanbozyel]. (2020, July 27). İstanbul Sözleşmesi Yaşatır! [Instagram photo]. Retrieved from <https://www.instagram.com/p/CDHuSnYpXVL/>
- Brechenmacher, S., & Brechenmacher, S. (n.d.). *Why Are Governments Weakening Protections Against Domestic Violence?* Carnegie Endowment for International Peace. Retrieved September 5, 2021, from <https://carnegieendowment.org/2020/09/01/why-are-governments-weakening-protections-against-domestic-violence-pub-82632>
- Bruns, A., & Highfield, T. (2015). Is Habermas on Twitter?: Social media and the public sphere. In *The Routledge companion to social media and politics* (pp. 56-73). Routledge.
- By News from Elsewhere. . . (2020, July 30). "Black and White" campaign highlights femicide in Turkey. Retrieved September 15, 2021, from <https://www.bbc.com/news/blogs-news-from-elsewhere-53596483>
- Cabrera, N. L., Matias, C. E., & Montoya, R. (2017). Activism or slacktivism? The potential and pitfalls of social media in contemporary student activism. *Journal of Diversity in Higher Education, 10*(4), 400.
- Calhoun, C. J. (Ed.). (1992). *Habermas and the public sphere*. MIT press.
- Chang, Hsia-Ching; Iyer, Hemalata (2012). "Trends in Twitter - Hashtag Applications: Design Features for Value-Added Dimensions to Future Library Catalogues". *Library Trends, 61* (1): 248–258.
- CHNGE™. [@chnge]. (2020, July 30). SWiPE  LEARN  SHARE 
[Instagram photo]. Retrieved from https://www.instagram.com/p/CDPj_fXgziF/
- Cwynar-Horta, J. (2016). The commodification of the body positive movement on Instagram. *Stream: Culture/Politics/Technology, 8*(2), 36-56.
- Crepax, R. (2020). The aestheticisation of feminism: A case study of feminist Instagram aesthetics. *ZoneModa Journal, 10*(1S), 71-81.
- Dailey, J. D. (n.d.). Honor killing | sociology. Retrieved September 3, 2021, from <https://www.britannica.com/topic/honor-killing>
- Dones Juristes. [@donesjuristes]. (2021, July 2). All our support to the Turkish companies in the face of the withdrawal of the Istanbul Convention from Turkey [Instagram photo]. Retrieved from <https://www.instagram.com/p/CQ0tqQTL-GT/>

- DuVernay, A. [@ava]. (2020, July 28). *Pretty sure this is just a cool excuse to post pix* [Instagram photo]. Retrieved from <https://www.instagram.com/p/CDJxOg9l7iP/>
- Eler, A. (2020, August 8). For Turkish Women Artists and Advocates, #challengeaccepted Is About More Than Just Selfies. Retrieved September 8, 2021, from <https://hyperallergic.com/581546/turkish-women-respond-to-challenge-accepted/>
- Goker, Z. G. (2019). Memories, stories and deliberation: Digital sisterhood on feminist websites in Turkey. *European Journal of Women's Studies*, 26(3), 313-328.
- Gómez, M. E. [@melgomez2015]. (2021, July 10). #istanbulconventionsaveslives [Instagram photo]. Retrieved from https://www.instagram.com/p/CRH5_Y-MRqf/
- Goñi, U. (2020, October 15). Argentina's women joined across South America in marches against violence. Retrieved September 18, 2021, from <https://www.theguardian.com/world/2016/oct/20/argentina-women-south-america-marches-violence-ni-una-menos>
- Habermas, J. (1991): "The public sphere" In Mukerji, C.; Schudson, M. (Ed.): Rethinking popular culture. Contemporary perspectives in cultural studies. Berkeley/Los Angeles: University of California Press. pp.398-404.
- Jackson, S.J. & Banaszczyk, S., 2016. Digital standpoints: Debating gendered violence and racial exclusions in the feminist counterpublic. *Journal of Communication Inquiry*, 40(4)
- Jamal, A. (2020, August 17). #ChallengeAccepted: Why the Instagram hashtag went viral, may be 'self promotional' and no, it did not originate in Turkey. Retrieved September 8, 2021, from <https://www.hindustantimes.com/fashion-and-trends/challenge-accepted-why-the-black-and-white-instagram-hashtag-went-viral-may-be-self-promotional-and-no-it-did-not-originate-in-turkey/story-cZtA0ycYdYMBYqNRXUoI0H.html>
- Kardashian, K. [@khloekardashian]. (2020, July 27).  Challenge Accepted  [Instagram photo]. Retrieved from <https://www.instagram.com/p/CDH3xn1BB3W/?hl=en>
- Lorenz, T. (2020, July 29). "Challenge Accepted" on Instagram: Black and White Selfies for Women. Retrieved August 29, 2021, from <https://www.nytimes.com/2020/07/27/style/challenge-accepted-instagram.html>
- Mahoney, J. (2015). The ice bucket challenge and its name calling strategy through social media: A descriptive statistical analysis. University of Louisiana at Lafayette.
- Mogensen, K. (2014). Henry Jenkins, Sam Ford & Joshua Green: Spreadable media: Creating Value and meaning in A Networked Culture, New York: New York University Press. 2013. *MedieKultur: Journal of Media and Communication Research*, 30(56), 4. <https://doi.org/10.7146/mediekultur.v30i56.16054>
- Moss, R. (2020, July 29). What Is #ChallengeAccepted On Instagram About? Here's The Story Behind It. Retrieved September 8, 2021, from https://www.huffingtonpost.co.uk/entry/this-is-the-whole-story-behind-the-challengeaccepted-black-and-white-photo-trend_uk_5f213145c5b66859f1f3e987

- Nabil, S. The Selfie Phenomenon—Opportunities for Empowerment or Strategies to Preserve the Status Quo?
- Ogan, C. L., & Baş, Ö. (2020). Women's Rights and Gender Equality in Turkey| Use of Social Media in the Struggle Surrounding Violence Against Turkish Women. *International Journal of Communication*, 14, 19.
- O'Malley, K. (2020, July 29). Challenge Accepted: The New Social Media Campaign Involving Black-And-White Selfies. Retrieved August 14, 2021, from <https://www.elle.com/uk/life-and-culture/culture/a33441480/challenge-accepted-instagram/>
- Özlem, A. (2021, June 2). Why Turkey's withdrawal from the Istanbul Convention is a global problem. Retrieved August 14, 2021, from <https://www.opendemocracy.net/en/can-europe-make-it/why-turkeys-withdrawal-from-the-istanbul-convention-is-a-global-problem/>
- ÖZŞENLER, D. (2021). Digital Activism: An Analysis of# challengeaccepted Hashtag Activism Developed for Violence against Women. *World Women Studies Journal*, 6(1), 35-62.
- Peltz, N. [@nicolaanpeltz]. (2020, July 28). *challenge accepted* [Instagram photo]. Retrieved from https://www.instagram.com/p/CDKJ_ANBkTf/
- Pressgrove, G., McKeever, B. W., & Jang, S. M. (2017). What is Contagious? Exploring why content goes viral on Twitter: A case study of the ALS Ice Bucket Challenge. *International Journal of Nonprofit and Voluntary Sector Marketing*, 23(1), e1586. <https://doi.org/10.1002/nvsm.1586>
- Pugh, F. [@florencepugh]. (2020, July 28). *Challenge accepted*. [Instagram photo]. Retrieved from <https://www.instagram.com/p/CDKRjbFl2Kc/>
- Renk. [@renk.magazin.]. (2021, July 7). *Austritt der Türkei* [Instagram photo]. Retrieved from <https://www.instagram.com/p/CRB2Rokgg0I/>
- Sebastian, M. (2019). Instagram and Gendered Surveillance: Ways of Seeing the Hashtag. *Surveillance & Society*, 17(1/2), 40–45. <https://doi.org/10.24908/ss.v17i1/2.12938>
- Şener, G. (2021). *Digital Feminist Activism in Turkey*. Media @LSE Working Paper Series No. 67. Retrieved from <https://www.lse.ac.uk/media-and-communications/assets/documents/research/working-paper-series/WP67.pdf>
- Sarıkaya, S. [@serenayss]. (2020, July 27). *#challengeaccepted* [Instagram photo]. Retrieved from <https://www.instagram.com/p/CDHuOzdpynh/>
- Smith, J. (2020, September 8). Students react to 'Challenge Accepted' trend aiming to raise Turkish femicides awareness. Retrieved September 3, 2021, from <https://thepolypost.com/arts-and-entertainment/2020/09/07/students-react-to-challenge-accepted-trend-aiming-to-raise-turkish-femicides-awareness/>
- Stewart, E., & Ghaffary, S. (2020, June 24). Political content has taken over Instagram thanks to Black Lives Matter. Retrieved September 8, 2021, from <https://www.vox.com/recode/2020/6/24/21300631/instagram-black-lives-matter-politics-blackout-tuesday>
- Thelwell, K. (2020, September 3). 5 Facts About Femicide in Turkey. Retrieved September 1, 2021, from <https://borgenproject.org/femicide-in-turkey/>

- Time.com. (2020, November 23). Co-Founder of ALS Ice Bucket Challenge, Which Raised \$200 Million for Research, Dies at 37. Retrieved September 8, 2021, from <https://www.dailyadvent.com/news/82ac8f529334d46d1ac14970b30a325c-CoFounder-of-ALS-Ice-Bucket-Challenge-Which-Raised-200-Million-for-Research-Dies-at-37>
- United Nations High Commissioner for Refugees. (n.d.-b). Gender-based Violence. Retrieved September 3, 2021, from <https://www.unhcr.org/gender-based-violence.html>
- Ustel, I. (2021). "Challenge Accepted" Movement on Instagram: An Embodied Virtual Protest. *Iowa Journal of Cultural Studies*, 21(1), 27-48.
- We Will Stop Femicide Platform. (n.d.). For English. Retrieved September 3, 2021, from <http://www.kadincinayetlerinidurduracagiz.net/for-english>
- Zhang, A. (2015). Women Uniting on the Web: Social Media Contributions against Gender-Based Violence in Turkey. Retrieved from <https://pioneeracademics.com/journals/> (p.427-461)



The Role of Iconoclasm in the History of Byzantine Art

Zehra Gundogdu

Author Background: *Zehra Gundogdu grew up in Turkey and attended Bahçeşehir High School for Science and Technology in Istanbul, Turkey. Her Pioneer research concentration was in the field of art history and titled “The Medieval Mediterranean: Confluence of Cultures.”*

Abstract

Byzantine Iconoclasm (726-843 CE) marks a transformational time in Byzantine history when the appropriateness of the icons used in religious settings was questioned on theological grounds. Extensive debates on the difference between icon veneration and adoration, and the icon itself with its archetype, brought forth various new statements and policies regarding the integration of icons in the renowned Byzantine monuments’ decorative programs. This paper explores the changes happening in the complexion of these monuments during the iconoclastic era and the way they reflected the complex course of the icon controversy using Hagia Sophia and the Chalke Gate of the Great Palace, Constantinople, as its principal case studies. In order to evaluate the correspondence between the diverse policies of the emperors and empresses of the era and the course of change in the art programs, a historical overview of the first era of Byzantine Iconoclasm (726-787 CE), the brief iconophile interlude following the period, and the Second Byzantine Iconoclasm Era (814-843 CE) is given. The art programs of the major Byzantine monuments of these eras are examined with a focus on the decorative elements subjected to iconoclastic activity. The changing display of the Christ Chalkites icon set on top of the Bronze Gate of the Great Palace, the debate surrounding the apse mosaics of Hagia Sophia depicting the Virgin and the Child, and the recurring replacement of figural mosaics with crosses in various churches such as the Church of the Koimesis in Iznik, Hagia Eirene in Istanbul, and the Hagia Sophia in Thessaloniki, testify to the dynamic change in Byzantine monuments reflecting the icon debate during Iconoclasm.

1. Introduction

Splendid, elaborate, and mystical, Byzantine art retains its value in the present day. Today, the Hagia Sophia in Istanbul, a monument that served as a center of religious, political, and artistic life for the Byzantine world for centuries, still leaves an unforgettable impression on visitors with its fascinating decorative mosaics program signifying the grandeur and mysteries of a distant era.¹ Sometime during Hagia Sophia's long history, a strong controversy making people question the appropriateness of these splendid figural mosaics and resulting in the removal of the artwork from its walls may seem most unexpected; however, this abrupt change of events in fact happened in 726-843 CE Byzantium. This era includes the first and second periods of Byzantine Iconoclasm, a time when the controversy over images took over the empire.

The icon controversy that arose with the rule of the Isaurian dynasty in 717 CE may be one of the most thought-provoking and transformative incidents of ideological change in Byzantine history. Located around the Mediterranean with its complex confluence of cultures, the Byzantine Empire underwent countless incidents of policy change due to changes in ideology throughout its history. However, the iconoclastic movement stands out as one of the most significant of these incidents.

Icons (from the Greek *eikones*) are defined as sacred images representing the saints, Christ, and the Virgin, as well as narrative scenes such as the Crucifixion (Figs. 1-3).² Iconoclastic debate centers on the appropriate use of these icons in religious veneration and is an extensive discussion that had been going on for centuries in Abrahamic religions. There were various arguments driving its course, but "fear that the viewer misdirects their veneration toward the image rather than to the holy person represented in the image lay at the heart of this controversy."³ "Iconoclasm" means "image breaking" and refers to a recurring historical impulse to destroy or alter images for various reasons.⁴

Starting from 726 CE, Byzantine Iconoclasm marks an era in which the evolution of religious ideology concerning figural art caused drastic change not only to the art program of Hagia Sophia, but also to many other renowned monuments of the era. Except for a fleeting iconodule resolution from 787 to 814 CE, the extensive debate around images continued until 843 CE, and the Byzantine Empire faced frequent changes in imperial policy regarding image

¹ "Hagia Sophia | History, Architecture, Mosaics, Facts, and Significance." *Encyclopedia Britannica*, Britannica, www.britannica.com/topic/Hagia-Sophia. Accessed 29 May 2021.

² Brooks, Sarah. "Icons and Iconoclasm in Byzantium." In *Heilbrunn Timeline of Art History*. New York: The Metropolitan Museum of Art, www.metmuseum.org/toah/hd/icon/hd_icon.htm. Published October 2001, Accessed 29 May 2021; Christ the Merciful (Christ Pantocrator) [Christus der Erbarmer; Christus der Barmherzige], 1st quarter 12th century. Artstor, library.artstor.org/asset/BERLIN_DB_10313748743; Icon with Gospel scenes originally lid of reliquary casket. 6th C. Artstor, library.artstor.org/asset/ARTSTOR_103_41822000458305; Medallion with Saint John the Evangelist from an Icon Frame. ca. 1100. Artstor, library.artstor.org/asset/SS7731421_7731421_11597359.

³ Brooks, Sarah. "Icons and Iconoclasm in Byzantium."

⁴ *Ibid.*

eneration throughout this history.⁵ Patriarchs resigned, the views of empresses and emperors clashed, clergymen were exiled, councils were summoned, the validity of their decrees were questioned, a debate equally philosophical and theological went on for years between the iconoclasts and iconophiles. The icon controversy has been discussed in such detail that the resulting Byzantine theorization of images exhibits an exceptional argumentative richness in terms of complexity and range.⁶

All these religious, political, and artistic statements have been made visible through the change to renowned monuments of the era. In such an environment, how did the monuments change over time to reflect the ongoing debates around the role of images? To understand the correlation between the course of the debate and the change that the monuments have undergone, we can consider an analogy: When inspecting the evolution of a person's mentality and trying to get an understanding of their changing thought process, one could look into the change in their habits and lifestyle. In most cases, the inner turmoils we face are reflected through tangible change in our lives. In a similar manner, when we look at communities, we see that with change in ideas through a natural process or through the intervention of authority, the abstract change projects its effects on the habits, lifestyle, and art of the community.

Since there is a new character to the renowned monuments of different eras evolving along with the character of the leading philosophical, sociological, and theological arguments, the transformation of arguments concerning images was reflected in the changing complexion of the art and architecture. As the figural imagery was analyzed for appropriateness on theological grounds, the varying opinions and policies of rulers on the icon debate designated the most current condition of figural imagery in the sites of veneration. The historical and theological contexts for the advent of iconoclasm in Byzantine Christianity will be explained to understand this correlation.

During periods of Byzantine Iconoclasm, figural representations faced interventions, elements of art programs were replaced, destroyed, and ultimately reinstalled. Along with the destruction, the performance of iconoclastic activity included transformations and alterations of figural art as well as the absence of images in newly built programs.⁷ The iconoclast movement also led to a fresh wave of production since in its aftermath the restoration of the destroyed parts of sites called for new design elements. Some important examples of post-iconoclasm Byzantine art refer back to the iconoclast period and celebrate its resolution. For example, reinstalled apse mosaics and a decorative program of tympana installed in Hagia Sophia in the ninth century include inscriptions against iconoclasm and depict iconodule patriarchs: a political statement on the victory over the iconoclasts.⁸

⁵ Apostolos-Cappadona, Diane, et al. "Iconoclasm." *Oxford Art Online*, 2003. *Crossref*, doi:10.1093/gao/9781884446054.article.t039783.

⁶ Alloa, Emmanuel. "Visual Studies in Byzantium: A Pictorial Turn Avant La Lettre." *Journal of Visual Culture*, vol. 12, no. 1, 2013, pp. 3–29. *Crossref*, doi:10.1177/1470412912468704.

⁷ Dudley, Lauren. "A Timeless Grammar of Iconoclasm." *Journal of Art Historiography* (11): 1, 2014, <https://arthistoriography.files.wordpress.com/2014/11/dudley-review.pdf>.

⁸ Natalia, Teteriatnikov. "Hagia Sophia, Constantinople: Religious Images and Their

These diverse results of the movement will be examined through an analysis of changes that the renowned Byzantine monuments underwent from 726 to 843 CE, the period of the Byzantine Iconoclasm. The monuments that will be analyzed in this context include Hagia Sophia in Istanbul (Fig. 4), the Chalke Gate of the Great Palace of Constantinople, the Church of the Koimesis in Iznik (Fig. 5), Hagia Eirene in Istanbul (Fig. 6), and Hagia Sophia in Thessaloniki (Fig. 7), with a focus on Hagia Sophia, Istanbul and the Chalke Gate.⁹

The continuous change in the display and product of art in Byzantine Iconoclasm almost directly reflects the discourse of the controversial icon debate. The effect of the theological movement on the renowned sites and art programs of the period visualized the dynamic nature of the debate itself.

2. Historical Overview

Christians debated the use of images almost from the beginning of their religion. The early Christian world's stance towards icons and religious images can be described as ambivalent. This ambivalence can be detected in early Patristic writings, some of which are based on a strong sense of the mysterious transcendence of God and the seductive danger of idolatry.¹⁰ From the early 4th century, the Christian community cultivated more concrete piety not necessarily centered around images, but centered on holy places, the tombs, relics of martyrs, and pilgrimage.¹¹ Veneration centered around these alternatives to religious imagery was predominantly considered "safer" with regard to the Old Testament prohibition "Thou shalt not make unto thyself any graven image, or any likeness of any thing that is in heaven above, or that is in the earth beneath" (Exodus 20:4). However, things began to change in the late 4th century. This period also coincides with the shift of the capital of the Roman Empire to Constantinople and the development of what ultimately became known as the Byzantine tradition.

Figural depictions first appeared in late 4th century Early Christian art, and became common in the 5th century. The illustrations of the period included representations of the Crucifixion and the Last Judgement.¹² Christian themes

Functional Context after Iconoclasm." *Zograf*, no. 30, 2005, pp. 9–19. *Crossref*, doi:10.2298/zog0530009t.

⁹ Hagia Eirene, Church of Holy Peace. Rebuilt 548, and after 740. Artstor, library.artstor.org/asset/AMCDONALDIG_10313947205; Mango, Cyril. "The Date of the Narthex Mosaics of the Church of the Dormition at Nicaea." *Dumbarton Oaks Papers*, vol. 13, 1959, p. 245. *Crossref*, doi:10.2307/1291137; Murat, Ali, Ali Murat, and Seda Sevensan. "Hagia Sophia to be open for all: Turkish president." *Digital Image, Anadolu Agency*, July, 2020; Zorin, Andrew. "Храм Святой Софии." *Digital Image*. July, 2011. <https://web.archive.org/web/20161019181928/http://www.panoramio.com/photo/58223010>.

¹⁰ Daley, SJ, Brian E. "God Visible." *Oxford Scholarship Online*, 2018, doi:10.1093/oso/9780199281336.001.0001.

¹¹ *Ibid.*

¹² Gelan, Cristina. "Ideology, Symbolism and Representation through Byzantine Art." *Anastasis (Iași)*, 2018, V (1): 130-145.

and figures were frequently depicted on household objects as well as sacred vessels and hangings, and the following centuries witnessed an explosion of Christian iconic creativity.¹³ Justinian's rebuilding of Hagia Sophia in the 6th century with its dazzling mosaics program testifies to this creativity. Support for the use of images is also displayed in a text from Leontios of Neopolis' book *Against the Jews*. Dating to the 630s or 640s, the text provides an example of a Christian narrative defending the appropriateness of icon veneration. It is written in the form of a dialogue between the Christian and the Jew, and the Christian is shown arguing in favor of the religious images.¹⁴

However, with the rise of the Isaurian dynasty in the 720s, the stance of the Byzantine world regarding icons changed dramatically.¹⁵ It has been pointed out by scholars that the decrees of the Quinisext Council (691-692) might have laid the foundation for this change of events.

Charles Barber has argued that Canon 82 of the council, "That artists are not to portray the Forerunner pointing to a lamb"¹⁶ - the first official statement about whether to depict Christ symbolically or anthropomorphically - lays the ground for iconoclasm since it highlights a need for a new type of religious representation.¹⁷

In his paper "The Medieval Art of Fear: Christ Pantokrator After Iconoclasm." Ravinder Binning also proposed that another decree of the same council has iconoclastic undertones. Canon 100 reads "That Those Things Which Incite Pleasures Are Not To Be Portrayed On Panels."¹⁸ After this decree, images in relation to sense-experience had to be reconsidered. Since images that incited sensations were deemed an irregularity, even a heresy, Canon 100 attacked images on theological grounds.¹⁹

Although these occurrences provided a potential baseline for the start of the icon controversy in Byzantium, the actions of Emperor Leo III Isaurikos (*reg* 717–41) in 726 CE mark the start of the controversial period of Byzantine Iconoclasm. He gave his support to a group of bishops condemning icons and ordered their destruction; following this change of policy, the patriarch of Constantinople resigned.²⁰

What lay within the roots of this doubt about the appropriateness of icon veneration is an extensive question resulting in many explanations. Most of the explanations are somehow connected to preventing a cardinal sin,

¹³ Daley, SJ, Brian E. "God Visible."

¹⁴ Barber, Charles. "The Truth in Painting: Iconoclasm and Identity in Early-Medieval Art." *Speculum*, vol. 72, no. 4, 1997, pp. 1019–36, doi:10.2307/2865956.

¹⁵ Daley, SJ, Brian E. "God Visible." *Oxford Scholarship Online*, 2018, doi:10.1093/oso/9780199281336.001.0001.

¹⁶ "Canon 82: Περὶ τοῦ μὴ τοὺς ζωγράφους ἐν τῷ δακτυλοδεικτεῖν τὸν Πρόδρομον ἀμὶνὸν ἐγχαράττειν." The Council in Trullo Revisited, ed. George Nedungatt and Michael Featherstone (Pontificio Istituto Orientale: Roma, 1995), 162.

¹⁷ Nelson, Robert S., and Charles Barber. "Figure and Likeness: On the Limits of Representation in Byzantine Iconoclasm." *The Art Bulletin*, vol. 85, no. 4, 2003, p. 797, doi:10.2307/3177371.

¹⁸ "Canon 100: Περὶ τοῦ μὴ ἐγχαράττεσθαι ἐν πίναξιν τὰ πρὸς ἡδονὴν κινουῦντα," Nedungatt and Featherstone, 180-181.

¹⁹ Binning, Ravinder S. *The Medieval Art of Fear: Christ Pantokrator After Iconoclasm*. ProQuest Dissertations Publishing, 2018.

²⁰ Temple, Richard. "Icon."

idolatry. As the icon veneration became increasingly popular, it became harder to regulate icon production and use. It was feared this cult of icons gave rise to inappropriate icon veneration.

Having said that what drew attention to this issue is a complex discussion, the start of the iconoclast movement in Byzantium corresponds with Islam expanding its domination over Asia Minor and Egypt. In the early years of Leo III's reign, the Muslim caliph Yazīd II (720-4) is recorded as demanding that Christian communities in his territory destroy mosaics and other religious images in churches.²¹ It is speculated that this iconoclastic policy of Muslims might have influenced Leo III to pursue similar action.

It is also important to consider a distinguishable feature of the period when searching for the reasoning behind iconoclast activity: glorification of the emperor's image. Because of the ban on icons in churches, artists focused on the execution of mosaics that highlighted the image of the emperor during the 8th-9th centuries.²² Whether this effect of the new policy against icons was a calculated one or not is unknown; however, it is undeniable that there was a political dimension to the emperor's intervention. As Peter Brown comments, "Iconoclasm was a tool to reaffirm the official and central religious authority of the Church over the holy man and his associate, the icon."²³

Following the reign of Leo III, Constantine V Kopronymos (*reg* 741–75) followed a similar, even more heavily iconoclastic policy. He summoned the Council of Hiaria, in which it was decided that icons separated or failed to represent Christ's dual divine and human nature. According to the views of the council, in "describing created flesh," the painter had either "circumscribed the uncircumscribable character of the Godhead [...]" or he had "confused the unconfused union [...], falling into the iniquity of confusion."²⁴ After the Council of Hiaria, iconoclasts started a series of violent actions against icons and iconophiles themselves, some of whom were even executed.²⁵

During the late 8th century, there was a lull in iconoclasm due to the efforts of the iconophile empress, Eirene (*reg* 797–802). She convened the Second Council of Nicaea (787), in which theological arguments in favour of icons were put forward and iconoclast policies were condemned.²⁶ She supported the restoration of the images destroyed and removed during the first wave of iconoclasm.²⁷ This iconophile dominance lasted until the reign of another iconoclast emperor, Leo V (*reg* 813-20).

²¹ Vasiliev, A. A. "The Iconoclastic Edict of the Caliph Yazid II, A. D. 721." *Dumbarton Oaks Papers*, vol. 9, 1956, p. 23, doi:10.2307/1291091, quoted in Daley, SJ, Brian E. "God Visible."

²² Gelan, Cristina. "Ideology, Symbolism and Representation through Byzantine Art."

²³ Brown, Peter. "A Dark-Age Crisis: Aspects of the Iconoclastic Controversy." *The English Historical Review* 88 (346), 1973, p. 1-34, quoted in Dudley, Lauren. "A Timeless Grammar of Iconoclasm."

²⁴ Elsner, Jas. "Iconoclasm as Discourse: From Antiquity to Byzantium." *The Art Bulletin* 94, no. 3, 2012, p. 368-394, doi:10.1080/00043079.2012.10786048.

²⁵ Temple, Richard. "Icon."

²⁶ Temple, Richard. "Icon."

²⁷ James, L. "Eirene." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000025720>.

The second iconoclastic period started with the abdication of the iconophile patriarch Nicephorus and the calling of a second iconoclast council in 815.²⁸ Although both the first iconoclasm era and this second wave are considered iconoclastic, they focus on different aspects of the icon debate. The assertions that images are not idols and their use is valid for didactic purposes were accepted in the second iconoclasm era. The focus of the era was preventing icon worship by selecting the contexts in which the use of icons was deemed appropriate. For example, icons were removed from positions near the ground to higher positions to decrease proximity with the viewer.²⁹ In the words of Jas Elsner, "The second iconoclastic era is not about images at all but about what counts as abuse in worship."³⁰ The actions of the era towards images were indeed iconoclastic; however, what drove them were different motivations.

The reign of Empress Theodora marked a reversal of iconoclasm as the empress was unsupportive of her husband Emperor Theophilus's (*reg* 829–42) iconoclastic policies. Exiles and prisoners were slowly released and icons, frescoes and mosaics were restored.³¹ The reinstatement of iconoclasm ended with the triumph of Orthodoxy at the Synod of Constantinople (843 CE), which was based on the formulations of icon defendant St John of Damascus.³² The Orthodox Church continues to celebrate this victory with an annual feast on the first Sunday in Lent.³³

This eventful history of debate surrounding the use of images in what becomes known as Byzantine art is exemplified in some ways by the history of decoration at renowned Byzantine monuments.

3. Case Studies

3.1 The Chalke Christ of the Great Palace

The principal residence of Byzantine emperors from Constantine the Great (*reg* 306–37) to Alexios I (*reg* 1081–1118), the Great Palace was the symbolic nerve center of the Byzantine empire.³⁴ It was a large complex of buildings, gardens, avenues, terraces, and private harbors situated on a site overlooking the Sea of Marmara to the southeast (Fig. 8).³⁵ Known as "the Chalke", its

²⁸ Elsner, Jas. "Iconoclasm as Discourse."

²⁹ *Ibid.*

³⁰ *Ibid.*

³¹ James, L. "Theodora, Empress of Byzantium." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000084458>.

³² Apostolos-Cappadona, Diane, et al. "Iconoclasm."

³³ Temple, Richard. "Icon."

³⁴ Magdalino, Paul, Walter B. Denny, Anthony Cutler, Robin Cormack, Lawrence E. Butler, Robert Ousterhout, Lioba Theis, Franz Rickert, Leslie Brubaker, Slobodan Curcic, and Hülya Tezcan. "Istanbul." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000042556>.

³⁵ "Istanbul, hypothetical plan of the Great Palace, 4th century AD to the late 15th: (a) Hipp...." *Grove Art Online*, Accessed 24 Jun. 2021, <https://www.oxfordartonline.com>.

Bronze Gate was the ceremonial gate of the inner palace. After Emperor Justinian's rebuilding of the palace in the 6th century, the Chalke's ceiling was decorated with mosaics depicting Justinian's victories and a portrait of the imperial couple surrounded by senators.³⁶

One of the most important icons of the period was an icon of Christ Chalkites shown standing on a footstool on top of this imperial gate, erected by Constantine the Great. The 6th century painting of the Christ Pantokrator in the Monastery of St. Catherine, Sinai (Fig. 9) is thought to be a copy of the Christ Chalkites on the Chalke Gate that hasn't survived to today.³⁷ There is a debate about its material and some scholars have argued that Chalkites was a painted wooden panel of Christ, as the iconoclasts destroyed it by fire.³⁸ However, in her paper "The Performative Icon", Bissera V. Pentcheva argues that identifying Chalkites as a metal relief icon depicting Christ is more accurate. Panel paintings are not suited for the exterior of a building, where humidity, light and temperature shift would ruin the tempera on wood, and indications of later sources point to a copper relief icon.³⁹

The first act of imperial iconoclasm was the removal of Christ Chalkites in 726 by Leo III Isaurikos. The icon on top of the Chalke gate was replaced with a cross. Since this act marked the start of Byzantine Iconoclasm, the Chalke has acquired an important position in the icon debate. In fact, the entire discourse on figurative images was thrown into question after the Chalkites' removal.⁴⁰ Many of the arguments of the period, whether they were for or against images, used Chalke as their debate focus. It is not an overestimation of importance to say that "its story summarizes the entire iconoclastic period."⁴¹

The image of Christ was restored in 787 by Empress Eirene, whose policies mark the end of the first Byzantine Iconoclasm era. However, with the reign of Leo V, history repeated itself. In 813, the icon was replaced with a cross yet again.⁴² Much like the first removal of the Chalkites Christ, its second removal marks the start of the second period of iconoclasm. In that aspect, the art program of the Chalke gate has directly reflected the course of the icon debate of the era.

Therefore, it was no surprise that after the final victory of Orthodoxy in 843, the Chalke gate, which acquired a symbolic meaning through the

com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-8000007254.

³⁶ Magdalino, Paul, Walter B. Denny, Anthony Cutler, Robin Cormack, Lawrence E. Butler, Robert Ousterhout, Lioba Theis, Franz Rickert, Leslie Brubaker, Slobodan Curcic, and Hülya Tezcan. "Istanbul." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000042556>.

³⁷ "Icon of Christ: Sinai, Monastery of St. Catherine: det.: after restoration.", https://library.artstor.org/asset/ARTSTOR_103_41822000458347; Ventura, Raphael, A. Dean McKenzie, and Susan Pinto Madigan. "Sinai." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000078884>.

³⁸ Pentcheva, Bissera V. "The Performative Icon." *The Art Bulletin* 88, no. 4, 2006, p. 631-55, <http://www.jstor.org/stable/25067280>.

³⁹ Ibid.

⁴⁰ Elsner, Jas. "Iconoclasm as Discourse."

⁴¹ Pentcheva, Bissera V. "The Performative Icon."

⁴² Magdalino, Paul, et al. "Istanbul."

controversy, was reunited with its Christ icon. The Chalke Christ was restored in a full-length standing version in mosaic, commissioned by Byzantine Empress Theodora (c. 810; *d* after 867).⁴³

3.2 Hagia Sophia

From the date of its dedication in 360 until 1453, Hagia Sophia served as the main cathedral of Constantinople. Following the rebuilding of the first church destroyed during riots in 404, the second Hagia Sophia was also largely destroyed by fire in the Nika riots of 532.⁴⁴ The present vast central-planned domed cathedral (Fig. 4), which replaced the second church on the site, was built and decorated by 537 during the reign of Justinian I (*reg* 527–65).⁴⁵

The decorative program of Hagia Sophia under Justinian's patronage mainly consisted of abstract designs of the marble revetment on the vertical walls combined with swathes of mosaic decoration including vegetal ornaments and crosses. Figurative decoration elements were also incorporated in the art program of the church under Justinian. Paulos Silentarios's verse ekphrasis recited early in 563 describes the silver images on the sanctuary screen of the church as showing Christ, the Virgin, and other saints.⁴⁶

Thus, prior to the first period of Byzantine Iconoclasm, icons and figural imagery were an established part of the decorative program of Hagia Sophia. However, as an important imperial monument at the heart of Constantinople, Hagia Sophia was too remarkable a Byzantine symbol to not be targeted by iconoclastic activity.

One of the ceremonial rooms of its patriarchal palace that was added to the southwest corner of the church (565-77) had a mosaic vault decoration including the bust of Christ and other saints. After the icon controversy took over Byzantium, this figurative decoration was cut out of the vault and destroyed by iconoclast Patriarch Niketas in 768.⁴⁷

There is also a controversy surrounding the apse mosaics depicting the Virgin and the Christ (Fig. 10).⁴⁸ The mosaics have been linked to a homily by Patriarch Phoitos delivered in Hagia Sophia on 29 March 867, and have been attributed to the reigns of the emperors Basil I (867-886) and Leo VI (886-912).⁴⁹ In contrast, Nicolas Oikonomides has proposed that the current image of the

⁴³ Alhas, Ali Murat, and Seda Sevensan. "Hagia Sophia to be open for all: Turkish president." Digital Image, Anadolu Agency, July, 2020; Elsner, Jas. "Iconoclasm as Discourse."

⁴⁴ Wegner, Emma. "Hagia Sophia, 532–37." *Heilbrunn Timeline of Art History*, New York: The Metropolitan Museum of Art, 2000, http://www.metmuseum.org/toah/hd/haso/hd_haso.htm.

⁴⁵ Magdalino, Paul, et al. "Istanbul."

⁴⁶ Mathews, Thomas F., and Cyril Mango. "The Art of the Byzantine Empire, 312–1453." *The Art Bulletin*, vol.56, no. 2, 1974, p. 279, doi:10.2307/3049237, quoted in Magdalino, Paul, et al. "Istanbul."

⁴⁷ Magdalino, Paul, et al. "Istanbul."

⁴⁸ "Madonna and child" mosaic from the apse, c. 834, Erich Lessing/ART RESOURCE, N.Y., Artstor, library.artstor.org/asset/LESSING_ART_1039901924

⁴⁹ Teteriatnikov, Natalia. "Hagia Sophia, Constantinople: Religious Images and their Functional Context After Iconoclasm." *Zograf : Časopis Za Srednjovekovnu Umetnost*, 2004-2005, (30): 9-19.

Virgin must have been first executed during the iconophile interlude of 787-815.⁵⁰

According to his argument, he was puzzled by certain differences regarding the position of the Virgin Mary between the description in Phoitos' homily and the actual mosaic in the apse. Phoitos described her as standing while the Virgin was depicted seated in the apse mosaic. Oikonomides concludes that the original mosaic of the seated Virgin was covered during the second iconoclastic period, and the Virgin Hodegetria described by Phoitos in 867 was executed on the new plaster. The original image was revealed during the earthquakes of 1343 and 1344.⁵¹

Having mentioned that some scholars argue for an earlier execution of apse mosaics in Hagia Sophia and propose that they were covered by iconoclasts, it is important to add that this argument is not the only effect of Iconoclasm on the said mosaic program. Hagia Sophia exemplifies the effects of the resolution of Iconoclasm on a monument as well as the effects of its progress.

The failure of the iconoclast movement to eradicate figural imagery from Byzantine art led to a massive celebratory redecoration in Hagia Sophia in the 9th century. Beginning from 867, the church received vast mosaic icons, and artists decorated the dome, semi-domes, apse and the tympanum above the nave. The first post-iconoclastic mosaics included the Virgin and the Child mosaics on the main apse and Archangel Michael & Archangel Gabriel in the arches in front of the apse (Fig. 11).⁵²

The new image of the Virgin and Child in the apse was accompanied by a contemporary inscription that is partially preserved at the north and south ends of the triumphal arch today.⁵³ The inscription reads, "The images which the impostors had formerly cast down here, pious emperors have again set up."⁵⁴ In addition to this political statement about the victory over Iconoclasm, the choice of images in the program of the nave also included elements of similar meaning. Along with the church's restoration after the earthquake of 869, images of Church Fathers were introduced to the decorative program (Figs. 12-13).⁵⁵ The inclusion

⁵⁰ Skhirtladze, Zaza. "The Image of the Virgin on the Sinai Hexptych and the Apse Mosaic of Hagia Sophia, Constantinople." *Dumbarton Oaks Papers* 68, 2014, p. 369-86, <http://www.jstor.org/stable/24643764>.

⁵¹ Oikonomides, N. "Some remarks on the apse mosaic of St. Sophia." *Dumbarton Oaks Papers* 39, 1985, p. 111-115, <https://doi.org/10.2307/1291518>.

⁵² Magdalino, Paul, et al. "Istanbul; Current Church was ordered by Emperor Justinian; designed by Greek scientists Isidore of Miletus and Anthemius of Tralles. Interior, Nave, Apse, Archangel Gabriel Mosaic. Current church built 532 - 537 AD; Mosaic probably date to the 9th century, set against the original 6th-century Golden Background, Image: April, 2010. Artstor, library.artstor.org/asset/ASITESPHOTOIG_10313842937.

⁵³ Teteriatnikov, Natalia. "Hagia Sophia, Constantinople"

⁵⁴ Mango, Cyril A. "Materials for the Study of the Mosaics of St. Sophia at Istanbul Vol. 8." Washington: Dumbarton Oaks Research Library and Collection, 1962.

⁵⁵ Isidore of Miletus and Anthemius of Tralles; dome rebuilt by Isidore the Younger; church converted into a mosque by Ottoman Sultan Mehmed II. Interior of Hagia Sophia, Upper Galleries Mosaics, St. Ignatius, St. John Chrysostom and St. Ignatius Theophorus. Current church 532 - 537 AD; the original 360 AD and the second 415 AD; converted into a Mosque circa 1453; Saint Mosaics date to the 9th AD. Artstor, library.artstor.org/asset/ASITESPHOTOIG_10313843108; Mango, Cyril, and Ernest J. W. Hawkins. "The Mosaics of St. Sophia at Istanbul: The Church Fathers in the North Tympanum" *Dumbarton Oaks Papers* 26, 1972, pp. 1, 3-41.

of the recent iconophile patriarchs Methodios and Ignatos, who were exiled during iconoclasm, highlighted the acceptance of their ideas and identified them as figures of major theological significance.

3.3 Other Sites with Evidence of Iconoclastic Activity

What we see at Hagia Sophia and the Chalke Gate of the Great Palace is symptomatic of similar activity found elsewhere in the Byzantine Empire during iconoclasm. Other examples include the Church of the Koimesis of the Theotokos (Fig. 5), Hagia Eirene in Istanbul (Fig. 6), Hagia Sophia in Thessaloniki (Fig. 7), Blachernae Palace and the Church of St. Mary of Blachernae, and the patriarchal palace in Constantinople.⁵⁶

During the iconoclastic periods, icons in churches were actively replaced with crosses, which were deemed the most appropriate decorative element for veneration. The figural representation at the Church of the Koimesis of the Theotokos, Iznik (Nikaia) shows traces of such intervention (Fig. 14).⁵⁷ The original seventh-century mosaics at the church depict Hetoimasia transforming into the Virgin holding the Child in the apse. This apse mosaic was replaced by the cross in the mid-eighth century.⁵⁸ The outline of this cross can be seen against the background of the apse mosaic. In fact, the mosaic was altered twice during iconoclasm. At first, iconoclasts only replaced the human figures with the cross, leaving the prepared throne, the hand of God, and the Trinitarian symbol of three rays that proceed from that hand.⁵⁹ This gradual alteration shows that as the dominant idea of what is appropriate concerning image veneration changed, the decorative program of the church was modified accordingly.

Similarly, the apse mosaic of the Hagia Eirene in Istanbul (Fig. 15) dating to the reign of Constantine V (*reg* 741–75) simply depicts a cross outlined in black on a pedestal against a gold ground.⁶⁰ During the same period, the Hagia Sophia in Thessaloniki, Greece was ornamented with a gold mosaic cross in the apse, as well.⁶¹ Traces of this cross can be seen against the gold background surrounding the present Virgin and Child mosaic that was set up after the end of iconoclasm (Fig. 16).⁶² The barrel vault of this sanctuary is also decorated with

⁵⁶ Hagia Eirene, Church of Holy Peace. Rebuilt 548, and after 740. Artstor, library.artstor.org/asset/AMCDONALDIG_10313947205; Mango, Cyril. "The Date of the Narthex Mosaics of the Church of the Dormition at Nicaea." *Dumbarton Oaks Papers* 13, 1959, pp. 245-52, doi:10.2307/1291137; Zorin, Andrew. "Храм Святой Софии." Digital Image, July, 2011, <https://web.archive.org/web/20161019181928/http://www.panoramio.com/photo/58223010>.

⁵⁷ Underwood, Paul A. "The Evidence of Restorations in the Sanctuary Mosaics of the Church of the Dormition at Nicaea." *Dumbarton Oaks Papers* 13, 1959, pp. 235-43, doi:10.2307/1291136.

⁵⁸ Pentcheva, Bissera V. "The Performative Icon."

⁵⁹ Nelson, Robert S. "Figure and Likeness: On the Limits of Representation in Byzantine Iconoclasm." *The Art Bulletin* 85, no. 4, 2003, pp. 797-800, doi:10.2307/3177371.

⁶⁰ Hagia Eirene, Church of Holy Peace. Rebuilt 548, and after 740. Artstor, library.artstor.org/asset/AMCDONALDIG_10313947198; Magdalino, Paul, et al. "Istanbul."

⁶¹ Daley, SJ, Brian E. "God Visible."

⁶² Anastasiadis, Aghis J., Sotirios Kissas, Georgios Velenis, and Robert Ousterhout.

aniconic figures dating back to the joint reign (780-797) of Constantine VI and Empress Eirene, when the iconoclastic controversy was still an ongoing debate, albeit not as strongly because of the empress' gradual efforts to resolve it. Segments of silver crosses and plant ornaments adorned the summit of the vault decorated with another multi-colored cross (Fig. 16).⁶³

Along with mosaics, series of frescoes, wood icons, and miniatures were destroyed since the iconoclastic movement condemned all kinds of figural religious representation. Frescoes of the patriarchal palace in Constantinople were painted over during the patriarchate of Nicetas (766-780), series of miniatures and wood icons of Blachernae Palace were destroyed.⁶⁴ It is also known that the Church of St. Mary of Blachernae, located at the same site as the Blachernae Palace, was a target of iconoclastic activity. The walls of the cruciform church were elaborately decorated with mosaics that iconoclast emperor Leo V altered.⁶⁵

The evolution of different kinds of decorative programs in these examples demonstrates the effect of the theological controversy on Byzantine art. These examples, along with many more, were the direct medium of the era in which the course of the icon controversy was reflected.

4. Conclusion

Byzantine Iconoclasm marked an era in history when a recurring theological debate amongst Abrahamic religions resurfaced and acquired complex layers with the intervention of leading political figures of the century. Extensive discussions on the use of images in religious settings, the difference between the icon and its archetype, and the appropriate form of icon veneration, brought about a continuous set of theological, political, artistic, and philosophical statements being made throughout the era. The direct subject of these varying ideas and policies was the product and display of Byzantine art. Considered the most artistically advanced monuments in Iconoclastic Byzantium, prominent examples of architecture studied in this paper played the role of a banner headline proclaiming the latest policy regarding icon veneration and set the standard. Decorative programs of these monuments were demonstrative of the continuous change of ideas in the unfolding icon debate in an almost directly reflective manner throughout iconoclasm.

As analyzed, starting from its official establishment with the reign of Leo III, Byzantine Iconoclasm gained a stronger impact, lost influence for a short period, resurfaced again, regained its significance, and was ultimately resolved

"Thessaloniki." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000084506>;

Salonika Haghia Sophia vault mosaic det. Theotokis, Hagia Sophia, 787-797 and 11th C, Artstor, library.artstor.org/asset/ARTSTOR_103_41822003253786.

⁶³ Anastassiadis, et al. "Thessaloniki."

⁶⁴ Daley, SJ, Brian E. "God Visible;" Gelan, Cristina. "Ideology, Symbolism and Representation through Byzantine Art."

⁶⁵ Dark, Ken, and Ferudun Özgümüş. *Constantinople: Archaeology of a Byzantine Megapolis*. Oxbow Books, 2013. *ProQuest Ebook Central*.

with the Triumph of Orthodoxy in 843 CE. This complex nature of the movement's history was observed in the change in art programs of various monuments of the era. As the debate proceeded, the Christ Chalkites icon set on top of the Chalke Gate of the Great Palace was the subject of many arguments regarding icons and was destroyed and set up twice by different emperors and empresses. Each change proclaimed to the public a new era regarding icon veneration: the first Byzantine Iconoclasm, its resolution, the Second Byzantine Iconoclasm, and ultimately the Triumph of Orthodoxy. In addition to the Chalke, the change of the mosaics program of Hagia Sophia's apse strongly reflected the course of the icon debate. Hagia Sophia also played a main role in the celebrations following the resolution of the controversy and was subjected to a major redecoration signifying the victory over iconoclasts. Many other buildings of the era followed a similar change in their art programs.

In the eyes of the modern viewer, these Byzantine symbols of grandeur that captivate visitors with their elaborate decoration might arouse the feeling of witnessing the everlasting state of the untouchable. The way popular culture reflects the controversies surrounding the use of images in religious art is also misleading because it generally associates iconoclastic activity exclusively with Islamic or Jewish settings. However, the changes that medieval Byzantine monuments underwent throughout Byzantine Iconoclasm provide attention-grabbing examples of icon debate transforming the complexion of artistic production. Be it sociological, political, or theological, the arts of the eras live and breathe the change happening in their community. Considering how reflective they were of the unfolding icon debate with their changing decorative programs, Byzantine monuments show that rather than being a glorious static view forever trapped in the time they were built, they embodied the change in imperial mindset throughout their history.

Bibliography

- Apostolos-Cappadona, Diane, et al. "Iconoclasm." *Oxford Art Online*, 2003. *Crossref*, doi:10.1093/gao/9781884446054.article.t039783.
- Alloa, Emmanuel. "Visual Studies in Byzantium: A Pictorial Turn Avant La Lettre." *Journal of Visual Culture*, vol. 12, no. 1, 2013, pp. 3–29. *Crossref*, doi:10.1177/1470412912468704.
- Anastasiadis, Aghis J., Sotirios Kissas, Georgios Velenis, and Robert Ousterhout. "Thessaloniki." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000084506>.
- Barber, Charles. "From Transformation to Desire: Art and Worship after Byzantine Iconoclasm." *The Art Bulletin*, vol. 75, 1993, pp. 7-16.
- Barber, Charles. "The Truth in Painting: Iconoclasm and Identity in Early-Medieval Art." *Speculum*, vol. 72, no. 4, 1997, pp. 1019–36, doi:10.2307/2865956.
- Barber, Charles. "The Truth in Painting: Iconoclasm and Identity in Early-Medieval Art." *Speculum*, vol. 72, no. 4, 1997, pp. 1019–36, doi:10.2307/2865956.
- Binning, Ravinder S. *The Medieval Art of Fear: Christ Pantokrator After Iconoclasm*. ProQuest Dissertations Publishing, 2018.

- Brooks, Sarah. "Icons and Iconoclasm in Byzantium." In *Heilbrunn Timeline of Art History*. New York: The Metropolitan Museum of Art, www.metmuseum.org/toah/hd/icon/hd_icon.htm. Published October 2001, Accessed 29 May 2021.
- Brown, Peter. "A Dark-Age Crisis: Aspects of the Iconoclastic Controversy." *The English Historical Review* 88 (346), 1973, pp. 1-34.
- Daley, SJ, Brian E. "God Visible." *Oxford Scholarship Online*, 2018, doi:10.1093/oso/9780199281336.001.0001.
- Dark, Ken, and Ferudun Özgümiş. *Constantinople: Archaeology of a Byzantine Megapolis*. Oxbow Books, 2013. *ProQuest Ebook Central*.
- Dudley, Lauren. "A Timeless Grammar of Iconoclasm." *Journal of Art Historiography* (11): 1, 2014, <https://arthistoriography.files.wordpress.com/2014/11/dudley-review.pdf>.
- Elsner, Jas. "Iconoclasm as Discourse: From Antiquity to Byzantium." *The Art Bulletin* 94, no. 3, 2012, pp. 368-394, doi:10.1080/00043079.2012.10786048.
- Gelan, Cristina. "Ideology, Symbolism and Representation through Byzantine Art." *Anastasis (Iaşi)*, 2018, V (1): 130-145.
- "Hagia Sophia: History, Architecture, Mosaics, Facts, and Significance." *Encyclopedia Britannica*, Britannica, www.britannica.com/topic/Hagia-Sophia. Accessed 29 May 2021.
- James, L. "Eirene." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000025720>.
- James, L. "Theodora, Empress of Byzantium." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000084458>.
- Kolrud, Kristine, Marina Prusac, and Anne Karahan. "Byzantine Iconoclasm: Ideology and Quest for Power" in *Iconoclasm from Antiquity to Modernity*, ROUTLEDGE, 2020.
- Magdalino, Paul, Walter B. Denny, Anthony Cutler, Robin Cormack, Lawrence E. Butler, Robert Ousterhout, Lioba Theis, Franz Rickert, Leslie Brubaker, Slobodan Curcic, and Hülya Tezcan. "Istanbul." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000042556>.
- Mango, Cyril. "The Date of the Narthex Mosaics of the Church of the Dormition at Nicaea." *Dumbarton Oaks Papers* 13, 1959, pp. 245-52, doi:10.2307/1291137.
- Mango, Cyril A. "Materials for the Study of the Mosaics of St. Sophia at Istanbul Vol. 8." Washington: Dumbarton Oaks Research Library and Collection, 1962.
- Mango, Cyril, and Ernest J. W. Hawkins. "The Mosaics of St. Sophia at Istanbul: The Church Fathers in the North Tympanum." *Dumbarton Oaks Papers* 26, 1972, pp. 1, 3-41.
- Mathews, Thomas F., and Cyril Mango. "The Art of the Byzantine Empire, 312-1453." *The Art Bulletin*, vol.56, no. 2, 1974, p. 279, doi:10.2307/3049237.
- Nelson, Robert S. "Figure and Likeness: On the Limits of Representation in Byzantine Iconoclasm." *The Art Bulletin* 85, no. 4, 2003, pp. 797-800, doi:10.2307/3177371.

- Oikonomides, N. "Some remarks on the apse mosaic of St. Sophia." *Dumbarton Oaks Papers* 39, 1985, p. 111-115, <https://doi.org/10.2307/1291518>.
- Pentcheva, Bissera V. "The Performative Icon." *The Art Bulletin* 88, no. 4, 2006, pp. 631-55, <http://www.jstor.org/stable/25067280>.
- Skhirtladze, Zaza. "The Image of the Virgin on the Sinai Hexptych and the Apsse Mosaic of Hagia Sophia, Constantinople." *Dumbarton Oaks Papers* 68, 2014, p. 369-86, <http://www.jstor.org/stable/24643764>.
- Temple, Richard. "Icon." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000039776>.
- Teteriatnikov, Natalia. "Hagia Sophia, Constantinople: Religious Images and their Functional Context After Iconoclasm." *Zograf: Časopis Za Srednjovekovnu Umetnost*, 2004-2005, (30): 9-19.
- Teteriatnikov, Natalia B. "The Mosaics of the Eastern Arch of Hagia Sophia in Constantinople: Program and Liturgy." *Gesta* 52 (1), 2013, pp. 61-84.
- Underwood, Paul A. "The Evidence of Restorations in the Sanctuary Mosaics of the Church of the Dormition at Nicaea." *Dumbarton Oaks Papers* 13, 1959, pp. 235-43, doi:10.2307/1291136.
- Vasiliev, A. A. "The Iconoclastic Edict of the Caliph Yazid II, A. D. 721." *Dumbarton Oaks Papers*, vol. 9, 1956, p. 23, doi:10.2307/1291091.
- Ventura, Raphael, A. Dean McKenzie, and Susan Pinto Madigan. "Sinai." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000078884>.
- Wegner, Emma. "Hagia Sophia, 532-37." *Heilbrunn Timeline of Art History*, New York: The Metropolitan Museum of Art, 2000, http://www.metmuseum.org/toah/hd/haso/hd_haso.htm.
- Zeitler, Barbara. "Theophilos, Emperor of Byzantium." *Grove Art Online*, 2003, <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000084478>.

Illustrations

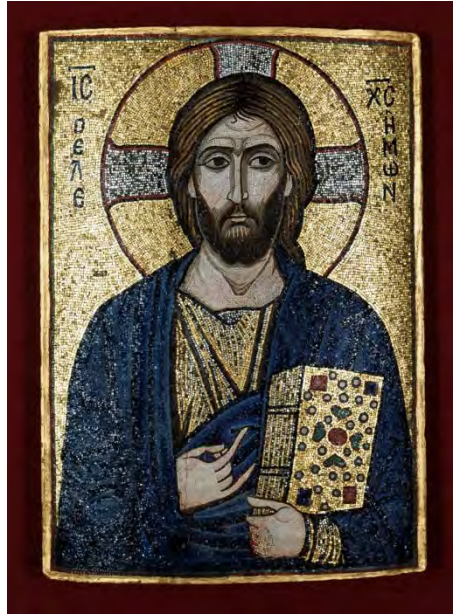


Figure 1. *Christ the Merciful Icon in Mosaic, 12th century CE (photographed by Jörg P. Anders, Berlin State Museums, Skulpturensammlung und Museum für Byzantinische Kunst, Staatliche Museen zu Berlin)*



Figure 2. *Icon with Gospel Scenes Originally Lid of Reliquary Casket, 6th century CE, Musée du Louvre (ARTstor Slide Gallery Collection, data from University of California, San Diego)*



Figure 3. *Medallion with Saint John the Evangelist from an Icon Frame Made in Constantinople, ca. 1100 (Work dedicated to public domain, The Metro-politan Museum of Art, Gift of J. Pierpont Morgan, 1917 17.190.67)*



Figure 4. *Hagia Sophia Exterior, rebuilt 6th century CE, General View, Istanbul (Anadolu Agency/ Getty Images)*



Figure 5. *Church of Koimesis in Iznik (Nicaea), 1870-1880s (Photo by Guillaume Berggren, Cyril Mango "The Date of the Narthex Mosaics of the Church of the Dormition at Nicaea", Dumbarton Oaks Papers)*



Figure 6. *Hagia Eirene, Istanbul, View from Southeast, Rebuilt 548, and after 740 CE (Publication: William L. MacDonald, Data and image from William L. MacDonald Collection, Princeton University, Department of Art and Archaeology)*



Figure 7. *Hagia Sophia Thessaloniki, Greece, built 7th century CE (By Andrew Zorin, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=53938343>)*

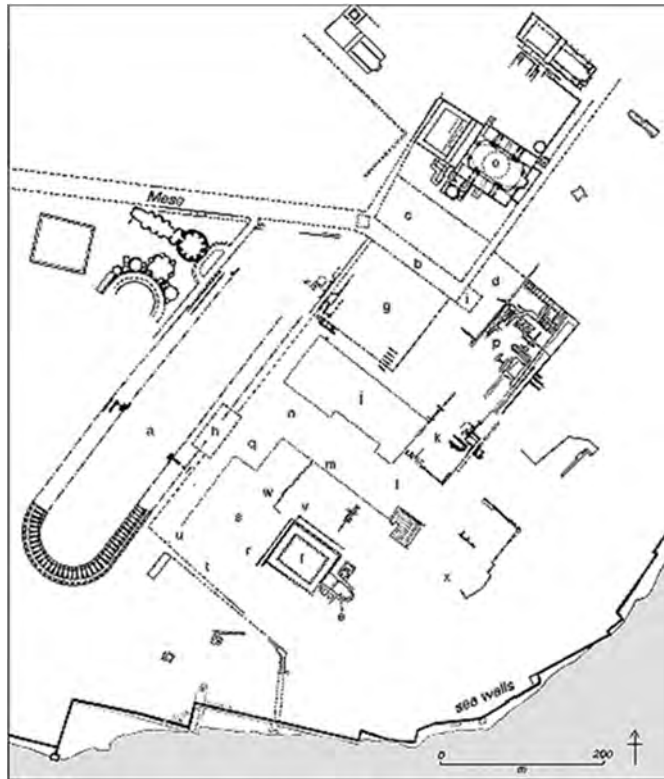


Figure 8. Istanbul, hypothetical plan of the Great Palace, 4th century AD to the late 15th: (a) Hippodrome; (b) Regia; (c) Augustaion; (d) Senate; (e) excavated apsed hall; (f) excavated peristyle court; (g) Baths of Zeuxippos; (h) Kathisma; (i) Chalke ('Bronze Gate'); (j) Delphax ('Tribunal'); (k) Triklinos of the Kandidates; (l) Consistorium; (m) Augusteus; (n) Triklinos of the 19 Couches; (o) Hagia Sophia; (p) Magnaura; (q) Daphne; (r) Chrysotriklinos; (s) Lausiakos; (t) Justinianos; (u) Skyla; (v) Trikonchos ('Hall of the Three Couches'); (w) Sigma; (x) Tzykanisterion

("Istanbul, hypothetical plan of the Great Palace, 4th century AD to the late 15th: (a) Hipp....." Grove Art Online. ; Accessed 24 Jun. 2021. <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-8000007254>.)



Figure 9. *Icon of Christ, Monastery of St. Catherine, Sinai, 6th century CE (ARTstor Slide Gallery, Data from University of California, San Diego)*



Figure 10. *Hagia Sophia Apse Mosaic, the Virgin and Child, 9th century CE (Image and original data provided by Erich Lessing Culture and Fine Arts Archives/ART RESOURCE, N.Y. <http://www.artres.com/c/htm/Home.aspx> ; <http://www.artres.com/c/htm/TreePflLight.aspx?ID=LES>)*



Figure 11. Archangel Gabriel Mosaic, Arch of the Apse, 9th century CE, Hagia Sophia (Image and original data provided by Shmuel Magal, Sites and Photos <http://www.sites-and-photos.com>)



Figure 12. Mosaics of Church Fathers, 9th century CE, Hagia Sophia, Istanbul (Image and original data provided by Shmuel Magal, Sites and Photos <http://www.sites-and-photos.com>)



Figure 13. *St. Ignatius the Younger Mosaic, late 9th century CE, Hagia Sophia, Istanbul (photo: © Dumbarton Oaks)*



Figure 14. *Apse Mosaics at the Church of the Koimesis (Dormition), 7th century CE, Iznik (Nikaia) (1912 Photo by N. K. Kluge)*



Figure 15. *Hagia Eirene Apse and Vault, 8th century CE, Istanbul (Publication: William L. MacDonald, Data and image from William L. MacDonald Collection, Princeton University, Department of Art and Archaeology)*



Figure 16. *Hagia Sophia Thessaloniki Apse and Vault Mosaics, 8th and 11th century CE, Greece (Artstor Slide Gallery Collection, Data from: University of California, San Diego)*



EDU Valuation: A New Era for Private Education or the Beginning of the End?

Zihao Gao

Author Background: *Zihao Gao grew up in China and currently attends McCallie School, in Chattanooga, Tennessee in the United States. His Pioneer research concentration was in the field of business and titled “Financial Accounting & Analysis, Capital Markets, Entrepreneurship, and Decision Making.”*

Abstract

New Oriental Education & Technology Group Inc. (EDU) is one of the most established private education providers in China. It focuses mainly on K-12 after-school tutoring, test preparation, and the educational materials distribution. The reputable brand and experienced management have fostered the company's sustainable growth, but the recent private education crackdown by Chinese government wiped out most of the ¥500 billion after-school tutoring market. To illustrate the crackdown's wide-reaching impact on the industry, the regulations are summarized and future prospects for EDU in uncharted territories, such as quality-oriented education and vocational training, are presented. Then, a two-scenario discounted cash flow method and a liquidation method are applied to verify whether the market is overly pessimistic. Based on the robust and conservative estimation, I conclude the market has been experiencing panic selling and recommend a target price of \$5.94 for EDU.

1. Introduction

Since 2000, the private education sector in China has exponentially grown to meet the surging demand of Chinese parents, who are eager to get their children as much tutoring as possible. Competition has intensified among young students to get into elite high schools and colleges, which can be achieved only by scoring high on the High School Entrance Exams (Zhongkao) and College Entrance Exams (Gaokao). From 2005 to 2020, the number of annual participants of Gaokao increased from 5.04 to 7.91 million, but the widely recognized colleges in China, namely “985” and “211 colleges,” accept only 0.8% and 2.4% of 19 million 18-year-old respectively each year.¹

¹ National Bureau of Statistics of China

About 63% of Chinese families have tutoring expenses accounting for more than 10% of the family's total expenses, just so their children will have a chance to succeed in entrance exams (Penguin Intelligence 2019).² Thus, instances where students take 8-10 hours of weekly classes outside school are not uncommon, accounting for two-thirds of their free time.

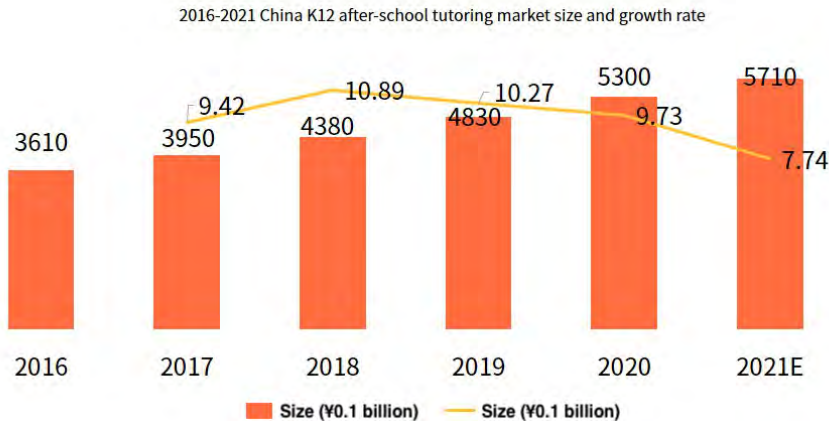


Figure 1: Chinese AST Market Size

Industry research conducted by iiMedia has estimated the tutoring market size to be ¥530 billion (\$81.8 billion), with compounded annual growth of 10.1% over the last five years, as shown in Figure 1 (iiMedia, 2020).³ One of the companies that benefited the most from this growing market is the New Oriental Education & Technology Group (EDU), one of the largest after-school tutoring providers, with a market capitalization over \$33 billion as of February 2021.

Founded in 1993, the New Oriental Education & Technology Group has the first mover advantage; they have diversified business operations, and their two largest competitors, TAL Education Group (TAL) and Gaotu Techedu (GOTU), were founded by former EDU employees. Experienced management and operations have carried the company through threatening setbacks, including recovery from Muddy Water Research's short sell (70% down) in 2012. In 2020, EDU had 10.6 million students enrolled, and the company has educated 64.9 million students since its founding.⁴ The total number of schools under the New Oriental Education & Technology Group has reached 1625, nearly 50% more than its closest competitor, TAL Education Group. The company had a consistent gross margin of around 55%, ten-years revenue CAGR at 24.9%, ten-year cash flow from operations CAGR at 12.4% and a return consistently higher than the S&P 500 benchmark for the past five years.

Though the industry looked promising, the Chinese government has cracked down on the private education sector, devastating future prospects for

² China For-profit Tutoring Market Consumer Purchasing Power Report, conducted by QQT and Tencent News on more than 20 thousand families in China having children between age 4-18: <http://www.199it.com/archives/971615.html>

³ IImedia Report | China K-12 Online Education Industry Research Report 2019-2020

⁴ <https://investor.neworiental.org/shareholder-services/investor-faqs/>

subset-relevant after-school tutoring.⁵ The 19th Meeting of Central Commission for Comprehensively Deepening Reform in May 2021 concluded the necessity of reducing unnecessary burdens for students in primary and secondary education.⁶ Tutoring for entrance exams of Mandatory Education will be banned on weekends, summer, and winter vacations, and existing subject-relevant tutoring institutions will be registered as non-profits and are restricted from foreign investment. This crackdown wiped out the half-a-trillion CNY market, and as the policy was gradually rolled out, more than 85% of EDU's market value evaporated year-to-date, resulting in a \$28 billion loss in value to shareholders. Figure 2 illustrates the price decline of the largest private education companies in China.

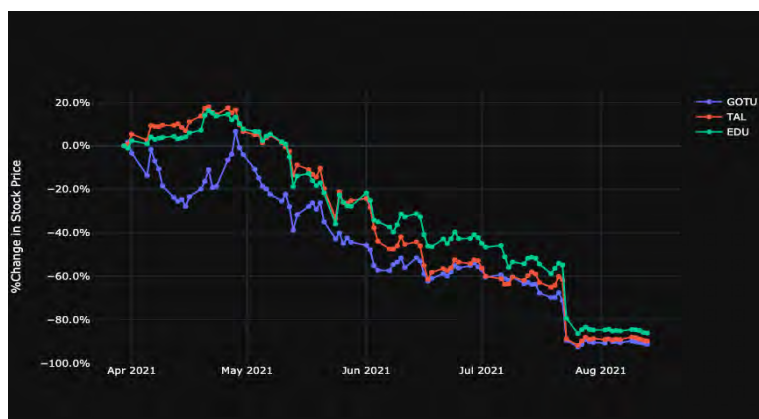


Figure 2: Stock price of GOTU, GSX, EDU %change

Although EDU has a strong brand name, ample cash reserves, an experienced management team, existing diversified operations, and new markets to explore. The market for vocational training was ¥281 billion in 2017, and is expected to grow 12.7% annually (L.E.K Consulting). And even with abundant resources to support the expansions, EDU still has a chance of being outcompeted by emerging companies like VIPidea or Iandcode in quality-oriented education, or established higher education providers such as China Education Group in vocational training.

This paper evaluates the company by analyzing the implications of policies and their impact on the recovery of the company's existing operations as well as potential expansion into vocational training. To value the company, I employ two valuation methods. The first is liquidation method and the second is a two-scenario discounted cash flow analysis, in which I provide a conservative and a more moderate forecast to determine whether the current expectations of

⁵ Subject-relevant after-school tutoring refers to tutoring in subjects included by the Entrance Exams

⁶ Document : *Opinions on Further Reducing the Burden of Homework and Off-campus Training for Students in Mandatory Education*, abbreviated as *Opinions*, the motives of the crackdown will be summarized in section: Industry Analysis

the company are reasonable. The lowest price supported by net liquidation value is \$0.97, and the implied share price is \$5.94 under the moderate estimation and \$3.4 under the pessimistic estimation, both of which are higher than the current \$2.35.⁷ Based on my analysis, I recommend a “strong buy” for the company.

2. Policy Analysis

The primary motivations behind the education crackdown are to comfort parents of their anxiety and to reduce burdens for young students.

The marketing approaches of tutoring companies have been outrageous for the past decade. At many private tutoring companies, teachers' KPI is directly linked to how many students they can retain (Economic Weekly 2021). Accounts of employees at such companies reveal how they are trained to ask and reply to questions in phone calls to exploit the anxiety of students' parents and persuade them to enroll their kids by exaggerating teachers' certifications and fictional success stories. Such phenomena grew worse overtime, and private institutions, such as TAL, GOTU, and Yuanfudao, have been investing heavily in promoting free-trial courses to lure customers in. As a result, the focus drifts away from teaching quality but to customer retention via misinformation.⁸

Therefore, the crackdown has strictly prohibited relevant marketing practices. Established business models with heavy emphasis on online education were not significantly harmed by the government policy's heavy regulation on marketing campaigns. However, smaller online competitors would likely to go out of business because they would have insufficient cashflow to support to pivoting in their offerings.

Additionally, the document places heavy emphasis on more efficient education in schools. A maximum of 90 minutes of homework can be assigned, and school-provided voluntary after-school activities are encouraged. Self-study hours are allowed in schools during weekday evenings, during which teachers could host Q&A sessions for students. If an employed school teacher is discovered to charge fees for extra-curricular tutoring, his/her PRC teaching certificate will be revoked.

However, the inequitable education quality among different schools makes schools unlikely to compete directly with established businesses capable of standardizing materials. If large businesses are to develop their quality-oriented education (QOE) segment, those who can afford it would still prefer their reputable services to schools. As such, schools would cater mostly to the needs of students with less financial support, which implies insubstantial competition between schools and large companies like EDU or TAL.

3. Company Overview

Founded in 1993 by Minhong Yu and located in Beijing, the New Oriental Education & Technology Group is one of the largest private education providers in China. The three primary sources of revenue in which EDU operates in are K-

⁷ As of September 1, 2021

⁸ According to in comprehensive estimates, the S&M expenses reach \$4.7 billion for the following five companies: EDU, TAL, GOTU, Yuanfudao, Zuoyebang.

12 after-school tutoring, test preparation, and books and other services. K-12 after school tutoring (AST) is generally subject-related and can be further broken down into AST for middle and high school students and AST for children. The test preparation includes overseas test preparation courses and PRC test preparation courses.⁹ Distribution of education materials accounts for the majority of books and other services, which consistently constituted 15% of the net revenue for the past five years. Approximately 260 titles and 17.1 million books authored or licensed by the company have been developed in China during the fiscal year ended in May 2021.

Courses are primarily delivered in classrooms for students in tier 1 cities, and for students in lower tier cities, a dual-teacher model is implemented: teachers' classes in top-tier cities broadcast through the OMO (Online Merge Offline) system and local assistant lecturers provide in-person guidance to students. This method alleviates the inequitable education resource distribution while making cost-efficient expansion feasible. In the Q3, 2020 earnings call, despite the current single-digit percentage revenue contribution by OMO system, the CFO expects this system to account for more income with slightly higher profit margin as it matures.

EDU's net income has been the highest among its competitors for the past decade except in 2019. Even during the pandemic when other competitors were heavily investing in online platforms and marketing, EDU's net income bounced to a historical high at ¥3 billion.¹⁰ This is attributable to EDU's caution in controlling the margin of each line item despite the fast growing revenue.¹¹ The historically sustainable development suggests the company is unlikely to approach a target aggressively and instills predictability into future forecast of the company's financial performance.

Even in the highly fragmented private education industry where the top companies' revenue account makes up less than 10% of the total market share, EDU still has a narrow economic moat due to outstanding academic performance and student enrollment in both tutoring and test preparation businesses. With the sprawling offline learning centers penetrating lower-tier cities, EDU is likely to benefit substantially from the rising average income in those emerging areas. These learning centers could potentially help the company pivot into QOE faster, as courses such as robotics, calligraphy, drawing, and musical instruments are better taught in-person.

Based on my policy analysis, I do not consider public schools' QOE programs as substantial competitors to those that EDU is developing. However, rising smaller competitors may intensify the competition, reducing the company's

⁹ Overseas: IELTS, TOEFL, SAT, SSAT, ACT, GRE, GMAT and LSAT;

PRC: CET 4, CET 6 and National Tests for Entrance into Master's Degree Programs.

¹⁰ During fiscal year 2020, TAL's net income was ¥-770M and GSX's net income was ¥267M

¹¹ COGS margin around 44%, sales & marketing margin 12%, and general & administrative margin 33%

QOE profit margin and ramping up sales and marketing expenses.¹² The increasing availability of test preparation materials may also diverge a portion of EDU's students, despite EDU's comprehensive services. In addition, the stagnant growth of test preparation, although partially attributable to the pandemic or the shifted focus on K-12 education, could suggest the declining quality. All these concerns will be elaborated in coordination with the industrial analysis in the following section.

4. Industry Analysis

With more students learning remotely and a comprehensive lock-down of major cities, companies dramatically increased marketing expenses to acquire more customers into their online learning platforms. The Covid pandemic in 2020 spurred the development of the online education industry. The remote-learning habit was cultivated on a massive scale under the comprehensive lockdown of major cities (Dhawan, S. 2020). Total enrollments in Koolearn (aka New Oriental Online) reached 285.2 million, a 30.89% increase YoY. Leading institutions have increased industry penetration by leveraging brand resource advantages. Established traditional private education providers, which operate offline, and online institutions, such as GOTU, are racing to acquire shares of China's online education market. The pandemic postponed operations of many offline institutions and effectively encouraged a large number of customers to go virtual.

However, online subsidiaries of well-established companies, such as Koolearn and Xuerersi Online (subsidiaries of EDU and TAL respectively) are very likely to successfully pivot into quality-oriented education along with offline businesses or adult education.¹³ EDU differentiates itself from other competitors in K-12 AST industries by highly diversified business lines, among which many have been explored for more than a decade¹⁴, offering a reliable base for further expansion.

Previously, TAL posed substantial threats as in recent years, Xuerersi Online had expanded to account for 29% of the mother company's (TAL's) growing revenue, whereas Koolearn only accounted for 3% of EDU's revenue in 2020. However, TAL focuses exclusively on K-12 education, the majority of which was banned by the mandate. In addition, one of Koolearn's major revenue sources is college education, accounting for 71.1%, 68.7%, and 59.38% of total revenue in 2018 to 2020 respectively with CAGR of 24%. Therefore, Xuersi would be less experienced in transitioning into higher education than EDU.

In addition to growing adult education in the online segment, another predictable and established line of business is test preparation. Test preparation is not targeted by regulatory policies in the education sector, and the demand has been consistent. Although the revenue attributed to test preparation is not disclosed, my conservative forecast indicates the percentage lies at roughly 20%

¹² Iandcode, positioned as a "mathematical thinking + coding", has raised \$10 million since its foundation in 2018.

VIPIdea, a rising unicorn providing English and art relevant QOE, has raised approximately \$1 billion in its financing rounds.

¹³ STEAM: Science, Technology, Engineering, Art, Mathematics

¹⁴ TOEFL preparation, GRE, Master Entrance Exam

to 30% of the total revenue for the last 4 years with CAGR of 7.5%.¹⁵ The number of enrollments in test preparation decreased from 612,000 in 2017 to 502,000 in 2020, though the average selling price has experienced substantial increases. Enrollments in overseas test preparation only experienced a 3% decrease compared to 4 years ago, but the cancellation of standardized tests such as SAT during the pandemic partially accounts for such decrease. The alarming decrease of enrollments in the PRC test could indicate a declining segment, but there could be a growing market for PRC test preparation.

Research has shown people in China have to become so overqualified for jobs that even undergraduates from top universities are unable to find a job with a salary that covers rent.¹⁶ As of 2019, among 557 million people with bachelors only 15.8% have pursued a master degree, and as job competition intensifies, this percentage is likely to have impactful increase. However, if EDU fails to stay competitive in PRC test preparation services, the segment's revenue may remain stagnant or continue experiencing slight decreases.

The outlook for overseas test preparation varies within an acceptable range from 280 thousand to 298 thousand. As first-tier cities continue to contribute international students to American and European countries and 36% of students enrolled in overseas test preparations come from Beijing or Shanghai, the segment future growth is predictable if not promising.

Beside test preparation, EDU has the potential to expand into vocational training. The number of public vocational training schools has increased, but there is a widening gap between the supply of highly trained technicians and the high demand by China's manufacturing industry. Five hundred thousand more well-trained technicians are still needed in the nation, and Table 1 shows the estimated gap of talent in the major ten sectors of manufacturing industry (Ministry of Education of PRC 2017).

Table 1: Official Estimates for Gap of Specialized Personnel

Specialized Personnel In Millions	2015	2020		2025	
		Total	Total (E)	Gap (E)	Total (E)
Top 5 Manufacturing Industries					
New Generation Information Technology	10.50	18.00	7.50	20.00	9.50
Electrical Equipment	8.22	12.33	4.11	17.31	9.09
New Generation Materials	6.00	9.00	3.00	10.00	4.00
CNC Machine Tools and Robotics	4.50	7.50	3.00	9.00	4.50
Marine Engineering Equipment	1.02	1.19	0.16	1.29	0.27

One policy providing long-term incentives to vocational training providers is the "1+X Certificate System".¹⁷ To address both college students'

¹⁵ Detailed procedure summarized in the Valuation section of this paper

¹⁶ Involution Literature Review and Research Prospects,

<https://www.sciscanpub.com/uploads/2021/01/04/pc20201110v3.pdf>

¹⁷ The system specifically targets: Modern Agriculture, Smart Manufacturing, High End Equipment, Next Generation Information Technology, Biomedicine, Energy Conservation, New Energy, New Materials, Digital Creativity, Modern Transportation, according to Implementation Plan of National Vocational Education Reform

employment problem (Zhang L. 2021), the country has been experimenting with the “1+X Certificate System” since March 2019. This system means students would earn multiple vocational skill-level certificates while earning a general certificate in their field of study.¹⁸ Although this certificate system may take longer than 8 years to be widely implemented, effective implementation would stimulate demand and benefit maturing vocational training providers.

I do not expect EDU to establish a mature system for training electrical engineers or high-end equipment operators in the forecast period, but there are many feasible areas in the short-term (5-8 years) which have existing certification tests and are scalable and standardizable. EDU could develop replicable curriculums to produce recognized results and maintain brand value.¹⁹ Compared with China Education Group, the largest higher education and vocational training provider in China, EDU still has an edge over enrollment.²⁰ Enrollment in EDU's PRC test preparation courses at the historical low in 2020 was already 213,000, whereas China Education Group had merely 250,300 students enrolled by February 2021. Nevertheless, China Education Group is a rising competitor, as its revenue tripled from 2016 to 2020, and is more experienced with vocational training services.

Based on the above analysis, I consider EDU to have a strong prospect if the company expands into both QOE and vocational training.

5. Valuation

5.1 Liquidation Method

Liquidation method is an asset-based valuation model typically used to determine whether the sale of a company's total assets would yield greater value than the net present value of the company's future cashflow. Applying liquidation method to a company with total debt less than one-third of cash and short-term investment is unusual, but liquidation method in this case is much more appropriate than multiple-based methods, such as P/S and EV/EBITDA, because historical sales and earnings information among private education providers are now inapplicable due to the crackdown on the business line that used to drive the growth of the industry.

GOTU and TAL are the only two competitors comparable in terms of market capitalization, so their liquidation value will be calculated and compared with that of EDU. The net liquidation value is given by:

Net Liquidation Value = Liquidation Value of Assets - Liquidation Value of Liabilities.

Liquidation Value of Assets:

Each line item is assigned a recovery rate, the expected liquidation value as a percentage of the book value. Both cash and short-term investments are expected to be fully recovered, as those liquid marketable securities mature in less than

¹⁸ The objective is to interface vocational skill-level standards and course syllabi and to maintain consistency between the objectives of training and the teaching materials.

¹⁹ e.g. IT: certificates approved by MSFT, IBM, Oracle... Accounting: ACCA, CPA ...

²⁰ Revenue of China Education Group in 2020 is ¥2.7 B whereas that of EDU is ¥25.5 B.

one-year. The two main items other than cash and short-term investments are prepaid-expenses and PP&E.

The companies' annual reports indicate prepaid expenses consist of prepaid rents, advances to suppliers, and advances to staff, all of which are difficult to retrieve, so the recovery rate for pre-paid expenses is 0 percent. EDU and TAL's top contributor for PP&E, accounting for at least 50% of total cost of PP&E is leasehold improvements, such as modification to existing structures, which will not be sold for a substantial percentage for their original cost. Construction in progress also won't be worth much nor would the depreciating buildings. The computer softwares could be sold for higher prices, but they do not account for much of the PP&E cost. These considerations lead to a conservative recovery rate of 35%. For each company, the recovery rates are set as follows:

Table 2: Expected Assets Recovery Rate

Cash And Equivalents	Short Term Investments	Accounts Receivable	Inventory	Prepaid Expenses	Other Current Assets	PP&E
100%	100%	75%	50%	0%	50%	35%

The liquidation value of asset i is then given by: $r_{recovery} \times Asset_{i,bookvalue}$ as shown in Table 2.

Table 3: Liquidation value of assets breakdown

CNY in millions	EDU	Liquidation Value (EDU)	TAL	Liquidation Value (TAL)	GOTU	Liquidation Value (GOTU)
Date	2021-02-28		2021-02-28		2021-03-31	
Cash And Equivalent s	10,162	10,162	2,911	20,993	20,993	2,911
Short Term Investments	29,186	29,186	2,449	17,443	17,443	2,449
Accounts Receivable, Net	52	39	0	126	168	0
Inventory	231	116	51	125	250	26
Prepaid Expenses	1,604	0	831	0	52	0
Other Current Assets	29	14	24	6,948	13,897	12
PP&E	16,438	5,753	1,645	4,661	13,317	576
Liquidation Value of Assets	41,456		43,699		5,579	

Net Liquidation Value:

The book value and the market value of liabilities are almost identical, so all liabilities must be subtracted. The net liquidation value of each company is as follows:

Table 4: Net Liquidation Value of Major Private Education Providers

CNY in millions	EDU	TAL	GOTU
Liquidation Value of Assets	41456	43699	5579
Total Liabilities	30,636	44,717	4,258
Net Liquidation Value	10820	-1018	1322

As of September 2, 2021, the number of outstanding shares for EDU is 1714 million, and the amount for GOTU is 256 million. The implied share prices supported by the net liquidation value for EDU and GOTU are \$0.97 and \$0.24 respectively using the exchange rate of ¥6.48 / \$1, which are much lower than EDU's five-year low at \$1.68 and GOTU's at \$2.25.

A negative liquidation value does not always suggest severe problems. TAL spent more than 1/3 of its revenue in sales and marketing, which partially fueled 32% growth in revenue from 2019 to 2020. In fact, TAL has the largest market capitalization since 2016 and will likely remain a formidable competitor of EDU due to its quality control, reputable brandname, and well-developed online system, as mentioned in the industry analysis.²¹

Such discrepancies reveal the suboptimal approximation without accounting for future prospects of sizable companies. In the following section, a two-scenario discounted cashflow analysis is performed and numerous aspects of EDU will be analyzed in detail to better reveal the fair value of the company.

5.2 Discounted Cashflow Method

The discounted cashflow method is a valuation procedure based on the net present value of an investment's expected future cashflow. Assumptions about estimated free cashflow (FCF) are made and the net present value of FCF is calculated. A terminal value is then calculated to account for cashflow beyond the projection period. Instead of using the traditional three-five split of historical and projected period forecast, I include 8 years of forecast. Major company restructuring takes time, so this timeframe will allow the pivoted business lines to stabilize.²²

The following procedures are performed in Discounted Cashflow Method: 1) Project FCF 2) Calculate WACC 3) Obtain terminal value 4) Calculate present value and determine valuation.

5.2.1 Projected FCF

Since the earnings release of fiscal year 2021 has been canceled, and at the time this paper was written fiscal year 2021 had passed, the projected period will not include 2021. The exclusion will not adversely affect the robustness of my

²¹ TAL has a quick ratio consistently over one for the past decade.

²² development of new curriculum, personnel recruitment, marketing campaign, renewed brand value...

valuation, because the crackdown policy was enacted near the end of EDU's fiscal year, therefore having a negligible impact to the company's operation in 2021. The general assumptions in each estimation are conservative; given the dire evaporation of market value of China concept stocks, an optimistic outlook is the last thing investors need. Two scenarios, one moderate and one pessimistic, are considered in the valuation process.

The moderate scenario is characterized by successful pivot into QOE, reasonable expansion of test preparation and vocational training, and a relatively higher perpetual cashflow growth rate. The pessimistic scenario is characterized by insufficient revenue recovery, stagnant growth of test preparation businesses, and a perpetual growth rate lower than GDP growth.

The two scenarios differ mainly in revenue projection. Except for revenue, the moderate estimation is shown while forecasting other income statement line items, because those items as a percentage of the revenue remain the same for both scenarios.

Revenue Estimation:²³

Historically, EDU's main revenue sources can be categorized into AST for secondary school students, AST for children, and books and other services. Since the Chinese government is unlikely to reverse its policies on tutoring for mandatory education, the original core of secondary school students can be discarded. Although the policy also banned subject-relevant tutoring for primary school and pre-school children, the subject-relevant tutoring is a much less important component. I expect this segment to be rebranded as QOE in the forecast period. EDU's wide offline learning center network allows the company to deploy relevant art, robotics, and music training much faster than its peers, as effective STEAM related education necessitates in-person learning (e.g. drawing, dancing, instruments, building robots, calligraphy). I expect in 8 years the company will be well capable of recovering from 40% to 130% of revenue in fiscal year 2020.

The declining enrollments for PRC test preparations are concerning, but overseas test preparations still have stable demand. As mentioned in Industry Analysis, I expect EDU to expand into vocational education, focusing on skill-level certification tests.²⁴ EDU's learning centers and developing OMO system would help with quality control and standardization of curriculum.²⁵ Short-term feasible areas include web/software development, accounting, and auditing. In the moderate estimation, I expect the combination of test preparation and vocational training to reach 170% of test preparation revenue in 2020. However, given the uncertainty of duration to establish a comprehensive course series and the potential scenario in which the company does not distinguish itself from established competitors, I expect this combination to reach 120% of test preparation revenue in 2020.

²³ Method for estimating historical revenue breakdown in Exhibit 1

²⁴ Industry Analysis section: EDU needs replicable curriculums and recognized results to maintain brand value.

²⁵ Quality Assurance Development (QAD): monitor & evaluate teachers' performance
Visible Progress System (VPS): standardize and maintain teaching quality

There is no detailed breakdown on customer profiles or sold books' type. The demand for tutoring books is not affected by the crackdown. The one part that would be gone from the revenue is the earning from books sold with a subject-relevant AST course.²⁶ I assume there will be a 30% decrease in book-related revenue as these customers leave. The expected target is 120% of FY2020's books and other services' revenue in the moderate estimation, and 90% in the pessimistic estimation.

Table 5: Estimated Revenue Breakdown

CNY in millions	Historical Period (E)				
	Year	2018	2019	2020	Restructure To
AST (mid/hi)		5,125	7,938	11,313	Spun off
AST for children		3,494	5,078	6,926	QOE
Test Preparation		4,324	4,279	4,829	Test Preparation & Vocational Training
Books and others		1,872	2,151	2,403	No Change

Table 6: Projected Revenue Breakdown 2022~2029 (Moderate)

CNY in millions	(Moderate) Projected Period								
	Year	2022	2023	2024	2025	2026	2027	2028	2029
QOE		2,770	4,848	6,233	7,156	8,080	8,542	8,849	9,003
Test Preparation & Vocational Training		5,070	6,117	6,814	7,279	7,744	7,977	8,132	8,209
Books & Others		1,923	2,283	2,523	2,684	2,844	2,924	2,977	3,004
Revenue		9,763	13,248	15,571	17,119	18,668	19,442	19,959	20,217
Revenue Growth		N/A	35.7%	17.5%	9.9%	9.0%	4.1%	2.7%	1.3%

Table 7: Projected Revenue Breakdown 2022~2029 (Pessimistic)

CNY in millions	(Pessimistic) Projected Period								
	Year	2022	2023	2024	2025	2026	2027	2028	2029
QOE		2,770	3,694	4,309	4,720	5,130	5,335	5,472	5,540
Test Preparation & Vocational Training		4,829	5,151	5,366	5,509	5,652	5,723	5,771	5,795
Books & Others		1,923	2,163	2,323	2,430	2,537	2,590	2,626	2,644
Revenue		9,522	11,008	11,998	12,658	13,319	13,649	13,869	13,979
Revenue Growth		N/A	15.6%	9.0%	5.5%	5.2%	2.5%	1.6%	0.8%

²⁶ Customers would previously pay for the course and get the books for "free", but in actuality a portion of the payment is counted as revenue from books on the income statement.

COGS, Sales & Marketing, SG&A:

EDU maintains COGS margin around 44%, S&M margin around 13%, G&A margin around 32 percent. The consistency offers a predictable long-term outlook. *Opinions* has indeed strictly regulated marketing expense of AST businesses but is not yet concerned with the marketing of QOE or adult and vocational education. As major competitors such as GOTU and TAL launch into a competition for market share in QOE, the short-term S&M margin is estimated to reach 16% from 13%. Eventually, as the pivoting businesses mature and the market share of EDU stabilizes near the end of the forecast period, I expect the S&M diminished to 11 percent.

Although EDU is relatively new to QOE and training new teachers will increase G&A expenses, the large-scale layoffs with the spin-off of K-12 subject-related AST decreases the size of operations to be managed. I expect the short-term general and administrative margin to stay around 35%. In the earning calls of Q2 2020, the OMO model is expected to reduce both COGS and SG&A. I expect COGS margin will be reduced slightly to 43.0% and G&A margin to 30 percent.

Table 8: Historical & Projected Income Statement Accounts (2018~2029)

CNY in millions	Historical Period			Projected Period								
	Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028	2029
COGS	6,832	9,503	11,339	4,296	5,810	6,806	7,459	8,107	8,415	8,610	8,693	
%rev	55.7%	60.6%	53.0%	44.0%	-	-	-	-	-	43.3%	43.1%	43.0%
S&M	2,079	2,654	3,178	1,562	2,009	2,238	2,333	2,411	2,380	2,316	2,224	
%rev.	14.0%	13.6%	12.5%	16.0%	15.2%	14.4%	13.6%	12.9%	12.2%	11.6%	11.0%	
G&A	5,093	7,140	8,175	3,417	4,536	5,215	5,609	5,983	6,095	6,121	6,065	
%rev	34.4%	36.7%	32.1%	35.0%	34.2%	33.5%	32.8%	32.0%	31.4%	30.7%	30.0%	

EBITDA: ²⁷

EBITDA's calculation is done by subtracting the cost of revenue and SG&A from revenue. During the past 6 years when the company was making sustainable expansions, the EBITDA margin stayed around 16%. Although much uncertainty remains, this historically consistent value could be indicative for the company's growth status. At the end of 2029, EBITDA margin reaches 16.0%, which suggests the company is restoring healthy growth and the last years' financial information is reliable for terminal value calculation.

Table 9: Historical & Projected EBITDA (2018~2029)

CNY in millions	Historical Period			Projected Period							
	Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028
EBITDA Margin	14.0%	13.4%	15.4%	5.0%	6.7%	8.4%	10.0%	11.6%	13.1%	14.6%	16.0%
EBITDA	2,192	2,872	3,925	488	893	1,311	1,719	2,167	2,551	2,911	3,235

²⁷ Earnings before Interest, Taxes, Depreciation, and Amortization

Depreciation and Amortization (D&A):

The D&A as a percentage of total revenue varied between 3-4% in the past decade and displays a growing trend for the past four years from 3.1% to 4.2 percent. This is attributable to the increased offline expansion: the total number of learning centers increased from 855 in 2017 to 1461 in 2020, a 70.9% increase. Nevertheless, the foreseeable drop in revenue and difficulty in raising capital will slow down the expansion speed dramatically, and consequently stabilize the rate of depreciation as shown below.

Table 10: Historical & Projected Depreciation and Amortization (2018~2029)

CNY in millions	Historical Period			Projected Period							
Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028	2029
D&A Margin	3.2%	3.7%	4.2%	4.2%	4.2%	4.2%	4.1%	4.1%	4.1%	4.0%	4.0%
D&A	507	787	1,078	412	555	647	706	764	790	804	809

EBIT:

Earnings before Interest and Tax is obtained by subtracting D&A expenses from EBITDA.

Table 11: Historical & Projected EBITDA (2018~2029)

CNY in millions	Historical Period			Projected Period							
Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028	2029
EBIT Margin %	10.7%	9.7%	11.1%	0.8%	2.6%	4.3%	5.9%	7.5%	9.1%	10.6%	12.0%
EBIT	1,686	2,085	2,847	76	338	664	1,013	1,403	1,762	2,107	2,426

Tax:

Given the uncertainty of company-specific taxation policies in the education sector, I assume a flat 25% effective tax rate (ETR).

Table 12: Historical & Projected Tax Expenses(2018~2029)

CNY in millions	Historical Period			Projected Period							
Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028	2029
ETR %	22.6%	28.4%	33.7%	25.0%	25.0%	25.0%	25.0%	25.0%	25.0%	25.0%	25.0%
Tax	381	592	959	19	85	166	253	351	440	527	607

Capital Expenditures (Capex):

To support initial expansion into QOE and vocational training, EDU is likely to increase spending on facilities, equipment, technology, and operating systems. As such, I expect capital expenditures to reach 9% of revenue in FY2020. Considering the existing wide-sprawling offline network EDU now has, I expect Capex to decrease gradually to 7% of revenue at the end of projected period.

Table 13: Historical & Projected Capital Expenditures (2018~2029)

CNY in millions	Historical Period			Projected Period							
Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028	2029
CapEx % sales	4.6%	8.7%	8.6%	9.0%	8.7%	8.4%	8.1%	7.8%	7.5%	7.3%	7.0%
CapEx	720	1,858	2,209	-879	-	-	-	-	-	-	-1,415
				1,150	1,304	1,383	1,455	1,462	1,448		

Change in Net Working Capital (ΔNWC):

When calculating the liquidity of a company, Net Working Capital is obtained by subtracting current liabilities from current assets. However, change in NWC used in FCF calculation excludes cash and cash equivalents and short-term investments. If change in cash and cash equivalents is included, subtracting change in NWC in the formula essentially subtracts part of the cashflow DCF is trying to estimate. Short-term investments are also not included in NWC due to lack of relevance to core operations and high unpredictability.

$$\text{FreeCashFlow} = \text{EBITDA} - \text{taxes} - \Delta \text{NWC} - \text{capitalexpenditures}$$

$$\text{NWC} = \text{CurrentAsset} - \text{CashCashEquivalents} - \text{ST-Investments} - \text{CurrentLiabilities}$$

Historically NWC as a percentage of revenue (NWC margin) has oscillated between -40% and -60%. Since there are no further details from the company, I assume NWC margin will stay at the middle of the historical range at -51% during the forecast period.

Table 14: Historical & Projected Change in Net Working Capital (2018~2029)

CNY in millions	Historical Period			Projected Period							
Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028	2029
ΔNWC moderate	-2,670	-	-1,295	-1,777	-1,777	-1,185	-790	-790	-395	-263	-132
ΔNWC pessimistic		2,188		-758	-758	-505	-337	-337	-168	-112	-56

Free Cashflow (FCF):

The fluctuation in FCF at the beginning of the forecast period is attributable to the uncertain change in NWC. Nevertheless, the converging FCF growth rate suggests such fluctuations do not impact the general valuation.

Table 15: Historical & Projected Free Cashflow (2018~2029)

CNY in millions	Historical Period			Projected Period							
Year	2018	2019	2020	2022	2023	2024	2025	2026	2027	2028	2029
FCF moderate	5,582	6,918	7,429	1,387	1,520	1,026	872	1,151	1,044	1,200	1,345
FCF Pessimistic				377	544	382	398	594	624	763	895

5.2.2 Calculate WACC:

The weighted average cost of capital (WACC) represents the cost of capital for a company. It is used to discount the future cashflows to their present values.

The general sentiment of foreign capital toward China has grown increasingly pessimistic after a series of regulatory crackdowns, and the official document assures that anti-trust regulations will remain in place.^{28 29} Therefore, the cost of raising capital is likely to go up, as shown in sensitivity analysis. In this valuation I use CAPM WACC, which contributes to a conservative estimate of expected returns for investors (Moore 2016).

$$WACC = (1 - ETR) \text{Cost of debt} \times \frac{D}{D + E} + \text{Cost of equity} \times \frac{E}{D + E}$$

The market value of the firm's equity, denoted by E, is given by:

$$\begin{aligned} \text{Equity} &= \text{No. shares outstanding} \times \text{share price} \\ &= 1714\text{M} \times \$2.35 = \$4.03\text{B} \\ \text{Equity} &= 6.48\text{¥}/\$ \times \$4.03\text{B} = \text{¥}26.1\text{B} \end{aligned}$$

Share price is the close price as of September 1, 2020, and the average exchange rate of USD to CNY in 2021 is 6.48 (X-RATES 2021). The share price uses the average taken during a period where continuous decline in share price occurred this year instead of the current market value (\$2.00), because the share price after panic selling distorts the market value of EDU's equity and consequently the firm's cost of capital.

Cost of Equity:

$$\text{Cost of Equity} = r_f + B \times (r_m - r_f)$$

(r_f is the risk-free rate of return, B is EDU's beta, and r_m is the expected return of the market)

Risk-free rate of return often uses US treasury bond rates, with the 20-year treasury rate quoted at 1.84% (Department of the Treasury). EDU's beta is estimated at 1.25 as of September 2021 (Tradingview, EDU). The NASDAQ Golden Dragon China Index (HXC) is used for r_m . Since the establishment of HXC until December 2020, the annual average rate of return was 10.8 percent.

$$\text{Cost of Equity} = 1.84\% + 1.25 \times (10.8\% - 1.84\%) = 13.0\%$$

²⁸ Timeline for Chinese government's crackdowns

Dec 2020, Ant Group IPO canceled

Apr 2021, Alibaba Group (BABA) fined for ¥2.8 billion (\$432 million)

May 2021, Tuition Guideline for after school tutoring

Jul 2021, Full Truck Alliance (YMM) app downloads suspended for investigation

Jul 2021, Didi (DIDI) app suspended for data privacy protocol violation

Jul 2021, Tencent Music ordered to terminate exclusivity agreements with musicians

²⁹ 2021-2025 Implementation Outline for Building Law-Based Government

Debt: EDU's total debt was ¥13.07 B at the end of Q3 2021, the latest quarterly report available.

Cost of Debt: Cost of Debt is obtained by total debt divided by interest expenses. In FY 2020, EDU had net interest expenses of ¥795.6 M and a total debt amounting to \$11276 M.

$$\text{Cost of Debt} = \frac{\text{¥795.6M}}{\text{¥13065M}} = 6.09\%$$

Effective Tax Rate: As mentioned in 4.2.1, I assume a flat ETR at 25%.

$$WACC = (1 - 25.0\%)6.09\% \times \frac{\text{¥13.07B}}{\text{¥13.07B} + \text{¥26.10B}} + 13.0\% \times \frac{\text{¥26.10B}}{\text{¥13.07B} + \text{¥26.10B}} = 10.21\%$$

5.2.3 Terminal Value (TV):

The terminal value accounts for all future cashflow after the projected period. I use the perpetual growth method instead of the exit multiples method. Since the private education industry focused mainly on subject relevant after-school tutoring prior to the crackdown, expectations on exit multiples such as EV/EBITDA are influenced substantially by the premise that subject-relevant after-school tutoring will fuel the growth of relevant companies. However, with this very business banned, historical EV/EBTDA values are no longer reliable.

In the moderate scenario, I expect the perpetual growth rate to be 8.2%, close to the convergence of the FCF growth rate in the forecast period, higher than the projected growth rate of China's GDP at 6 percent (Statista 2021). As the emerging lower tier cities to produce more high-income households and Chinese families' tendency to spend additional income on children's education, I assume FCF growth will outpace the economic growth in long-term. This perpetual growth rate has an implied EBIDA exit multiple of 22.4, and a EV/EBIDA of 11.9. This parameter is conservative for a company with a ten-year cash flow from operations CAGR of 12.4 percent. In the pessimistic scenario, the growth rate is expected to be 1% lower than the GDP growth rate, since in this situation the company would no longer be competitive enough to maintain strong growth. This perpetual growth rate produces an implied EIBTDA exit multiple at 8.4.

$$\text{Terminal Value} = \frac{FCF_n(1 + r_g)}{WACC - r_g}$$

(FCF_n : free cash flow at the end of the forecast period; r_g : the perpetual growth rate of FCF)

$$TV_{\text{moderate}} = \frac{1345(1 + 8.5\%)}{10.21\% - 8.5\%} = \text{¥72764M} \quad TV_{\text{pessimistic}} = \frac{895(1 + 5.2\%)}{10.21\% - 5.2\%} = \text{¥18831M}$$

5.2.4 Present value and determine valuation

Discounted Free Cash Flow:

$$\text{Moderate: Discounted FCF} = \sum_{n=1}^{n=8} \frac{FCF_n}{(1+WACC)^n} = ¥6382M$$

$$\text{Pessimistic: Discounted FCF} = \sum_{n=1}^{n=8} \frac{FCF_n}{(1+WACC)^n} = ¥2856M$$

Table 16: Historical & Projected Discount Factors (2018~2029)

Year	2022	2023	2024	2025	2026	2027	2028	2029
Discount Factor	90.7%	82.3%	74.7%	67.8%	61.5%	55.8%	50.6%	45.9%
Discounted FCF (moderate)	342	448	286	270	366	348	386	411
Discounted FCF (pessimistic)	1,258	1,251	766	591	708	582	607	618

Discounted Terminal Value (DTV) calculation:

$$\begin{aligned} DTV_{\text{moderate}} &= \frac{TV_{\text{moderate}}}{(1+WACC)^8} = \frac{¥72383M}{(1+10.21\%)^8} = ¥33256M \\ &= \frac{TV_{\text{pessimistic}}}{(1+WACC)^8} = \frac{¥18790M}{(1+10.21\%)^8} = ¥8633M \end{aligned}$$

Implied equity value: The sum of discounted future free cash flow, discounted terminal value, cash and cash equivalents, and short-term investments subtracting total debt yields the implied equity value. Short-term investments consist of mainly of financial products purchased from Chinese banks and trusts and have maturities ranging from one month to less than one year.³⁰ As such, short-term investments could be considered as equivalents to cash and are added to the implied equity value.

$$\text{ImpliedEquityValue} = \text{DiscountedTerminalValue} + \text{DiscountedFCF} + \text{Cash} + \text{ST-Investment} - \text{Debt}$$

Moderate:

$$¥6382M + ¥33256M + ¥10162M + ¥29185M - ¥13065M = ¥65920M$$

Pessimistic:

$$¥2856M + ¥8633M + ¥10162M + ¥29185M - ¥13065M = ¥37771M$$

Implied share price:

$$\begin{aligned} \text{Implied Share Price} &= \text{Implied Equity Value} / \text{SharesOutstanding} \\ \text{SharePrice}_{\text{moderate}} &= ¥65920M / 1714M \text{ shares} = ¥38.46 \times \$1 / ¥6.48 = \$5.94 \\ \text{SharePrice}_{\text{pessimistic}} &= ¥37771M / 1714M \text{ shares} = ¥22.04 \times \$1 / ¥6.48 = \$3.40 \end{aligned}$$

5.2.5 Sensitivity Analysis:

Factors such as uncertain outlooks, regulatory risks, and market sentiment can vary the cost of capital and EBITDA exit multiples. To provide a range for

³⁰ Annual Report, F-32

fluctuations due to these factors and to examine the robustness of DCF Analysis, a sensitivity analysis is performed on both the moderate and the pessimistic scenarios. The foreign exchange rate used is \$1 / ¥6.48.

Table 17: Sensitivity Analysis of Moderate Scenario ($r_g = 8.2\%$)

CNY	Exit Multiple (EBITDA)					USD	Exit Multiple (EBITDA)				
WACC	20.38	22.38	24.38	26.38	28.38	WACC	20.38	22.38	24.38	26.38	28.38
8.21%	39.51	41.52	43.53	45.54	47.54	8.21%	6.10	6.41	6.72	7.03	7.34
9.21%	38.06	39.93	41.79	43.66	45.52	9.21%	5.87	6.16	6.45	6.74	7.03
10.21%	36.73	38.46	40.19	41.93	43.66	10.21%	5.67	5.94	6.20	6.47	6.74
11.21%	35.49	37.11	38.72	40.33	41.95	11.21%	5.48	5.73	5.98	6.22	6.47
12.21%	34.36	35.86	37.36	38.86	40.37	12.21%	5.30	5.53	5.77	6.00	6.23

Table 18: Sensitivity Analysis of Pessimistic Scenario ($r_g = 5.2\%$)

CNY	Exit Multiple (EBITDA)					USD	Exit Multiple (EBITDA)				
WACC	6.40	8.40	10.40	12.40	14.40	WACC	6.40	8.40	10.40	12.40	14.40
8.21%	21.44	22.83	24.22	25.61	27.00	8.21%	3.31	3.52	3.74	3.95	4.17
9.21%	21.13	22.42	23.71	25.00	26.29	9.21%	3.26	3.46	3.66	3.86	4.06
10.21%	20.84	22.04	23.24	24.44	25.63	10.21%	3.22	3.40	3.59	3.77	3.96
11.21%	20.57	21.69	22.80	23.92	25.03	11.21%	3.17	3.35	3.52	3.69	3.86
12.21%	20.32	21.36	22.40	23.44	24.48	12.21%	3.14	3.30	3.46	3.62	3.78

6. Methods Evaluation & Conclusion

Given the high uncertainty involved in EDU's future, different valuation methods ought to be carefully evaluated to make a sensible investment recommendation. The liquidation method produces an implied share price target at \$0.97, whereas the DCF analysis implies a price target at \$3.40 in the pessimistic scenario and at \$5.94 in the moderate scenario.

The liquidation method essentially produces a floor for EDU's valuation. The success of new product lines is far from guaranteed, and the possibility of the Chinese government further cracking down the education industry, targeting areas like overseas study training and further eroding EDU's business, cannot be neglected. In the scenario in which EDU would have to liquidate now, the implied target price at \$0.97 covers 42% of the closing price on September 1, 2021, a remarkable percentage for a risky company, and also the highest among its competitors.

The DCF analysis, on the other hand, offers a comprehensive forecast into EDU's future operation status, and the two scenarios explore the cases in which EDU successfully makes a pivot to new product areas like vocational training and QOE, and which EDU fails to capture market shares into the new areas, respectively. The DCF analysis for the moderate scenario is performed under the estimation that FY2022's revenue is cut by 60% and achieves a revenue CAGR of 10.1% by the end of FY2029 for a company with a 24.9% historical revenue CAGR. Along with a conservative perpetual growth rate of FCF, the analysis yields \$5.94, representing a 157% upside to its current stock price. DCF analysis for the pessimistic scenario is carried out under the unlikely situation where all of EDU's operations become stagnant with a revenue CAGR less than 5.7 percent, yielding an implied price at \$3.40, representing a 47% upside.

Based on the implied price target at \$5.94 in the moderate scenario, \$3.40 in the pessimistic scenario, and \$0.97 using liquidation method, I believe

the upside is sufficiently large to justify the business risk EDU is undertaking in pivoting its business. Moreover, when considering my analysis of EDU's strong brand and operational experience, there is a strong possibility to success. Therefore, this paper recommends a "strong buy" for EDU at the current stock price target at \$2.31.

The industrial crackdown posed unprecedented challenge to EDU, and it is too rash to conclude this education giant would thrive in the unexplored territories with intense competition. In the fragmented education industry with uncertainties, it is even possible for companies with the amplest cash reserves to fall out of the race. The OMO system and extensive web of learning centers will be key to their sustainable growth, standardization of its curricula and brand reputation in both quality-oriented education and vocational training, and time will tell whether now is the beginning of a new era.

Exhibit 1: Revenue breakdown estimates

In the estimation process, I created skewed normal distributions of course prices based on the reported price range for the expected price per enrollment per course. The more negatively skewed distribution indicates a more negative difference between the mean and the median. The expected price was then multiplied by reported enrollments and estimated courses enrolled per paying customer. The skewed normal distribution with skewness α is given below:

$$f(x) = 2\phi(x)\Phi(\alpha x) = \int_{-\infty}^{\alpha x} \alpha \phi(t) dx = \int_{-\infty}^{\alpha x} \alpha \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dx$$

The (x_1, x_2) on the x axis is transformed into the price range (p_1, p_2) . The expected price is calculated by integrating the distribution given by $f(p)$ displayed as follows.

$$p = x_1 + (x - x_1) \frac{(p_2 - p_1)}{100}$$

$$E(p) = \int_{x_1}^{x_2} f(p) dx$$

After tuning parameters, revenues of AST for secondary school students, AST for children, and test preparation have α values -2, -1, and 0 respectively. $x_1 = -5$ and $x_2 = 5$, but $x_1' = -3$ to reduce the composition of low prices in the distribution, as they are rare in reality but inflate the expected value in a left-skewed distribution.

To reduce the impact of extremely high prices in test preparation, p_2 for after school tutoring is set in the middle between the lowest and the highest reported course prices.³¹ The link attached includes the detailed parameters and the python code with concise explanations.³² Table 3 compares the estimated revenue and reported revenue of 2018, 2019, and 2020. The consistency implies the robustness of the estimation.

³¹ the price range of test preparation ranges from ¥1500 to ¥49800 (\$231 - \$7685), a wide gap misleading to the estimation process

³² <https://drive.google.com/file/d/1H7JdN7P3Xd02Iv5v59TDrx7pII6Gm0Q-/view?usp=sharing>

Table 19: Estimated Revenue Accuracy

CNY in millions	2018-05-31	2019-05-31	2020-05-31
Reported Total Revenue	15,689	21,382	25,539
Estimated Total Revenue	14,815	19,446	25,471
AST (secondary)	5,125	7,938	11,313
AST (primary)	3,494	5,078	6,926
Test Preparation	4,324	4,279	4,829
Books and others	1,872	2,151	2,403
Estimated Revenue % Reported Revenue	94.4%	90.9%	99.7%

Works Cited

- Beijing Youth Daily. (2013, September 2). *The secret of the after-school tutoring: the teachers as salesmen*. people.cn. <http://edu.people.com.cn/n/2013/0902/c1053-22771678.html>.
- China GDP Growth Rate 2011-2024.” *Statista*, 12 Apr. 2021, www.statista.com/statistics/263616/gross-domestic-product-gdp-growth-rate-in-china/
- Daily Treasury Yield Curve Rates*, U.S. Department of the Treasury, 1 Sept. 2021, www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yield.
- Dhawan, S. (2020). Online learning: A panacea in the time of covid-19 crisis. *Journal of Educational Technology Systems*, 49(1), 5–22.
- Economic Weekly. (2021, March 18). Millions Salaries for Teacher Poaching and Teachers' Transformation to Salesmen. Phoenix New Media. <https://finance.ifeng.com/c/84hoTnasfTb>.
- EDU Stock Price and Chart - NYSE:EDU.” *TradingView*, 5 Sept. 2021, www.tradingview.com/symbols/NYSE-EDU/?exchange=NYSE.
- Hao, Yuran. The State Council of China Issued the Implementation Outline for the Construction of the Communist Party of China (2021-2025).”, Xinhua News Agency, 11 Aug. 2021, www.gov.cn/zhengce/2021-08/11/content_5630802.htm.
- Ministry of Education of People's Republic of China. *Manufacturing Talent Development Planning Guide*. 2017.
- Moore, David J. A Look at the Actual Cost of Capital of US Firms.” *Cogent Economics & Finance*, vol. 4, no. 1, 2016, doi:10.1080/23322039.2016.1233628.
- Penguin Intelligence , Tacent News. China For-Profit Tutoring Market Consumer Purchasing Power Report.” *119it*, 27 Nov. 2019, www.199it.com/archives/971615.html.
- Sharma, Anip. China Investment Outlook 2019 Spotlight on Edtech: L.E.K. Consulting.” L.E.K Consulting, 24 Feb. 2019, www.lek.com/insights/sr/2019-china-edtech-outlook.
- X-RATES. Monthly Average Converter Chinese Yuan Renminbi per 1 US Dollar Monthly Average.” *X-RATES*, 2021, www.x-rates.com/average/?from=USD&to=CNY&amount=1&year=2021.

- Xuan, Yu, et al. Involution Literature Review and Research Prospects.” *Psychology of China*, vol. 2, no. 12, 2020, doi:10.35534/pc.0212092.
- Zhang, L. (2021, August 19). China's multipronged approach to supporting job-seeking of college grads. http://www.xinhuanet.com/english/2021-08/19/c_1310135135.htm.



Migration's Effect on the Sending Country: The Case of Poland

Ziyu Zhou

Author Background: *Ziyu Zhou grew up in China and Poland and currently attends BASIS International School Park Lane Harbour in Huizhou, China. Her Pioneer research concentration was in the field of economics and titled "International Economics."*

Abstract

After the accession of Poland into the European Union in 2004, the country faced mass migration outflow. The phenomenon is collectively referred to as post-accession migration and has become one of the most remarkable contemporary migration movements in Europe. Such drastic changes in the Polish population are bound to have impacts on the nation's economy: not just direct impacts in the labor market, but also indirect impacts in other sectors of the economy. One characteristic of post-accession migration is a strong work-driven motive, shown by the origin, destination, and age of most Polish migrants. The aim of this paper is to explore the relationship between net migration and industry growth in the post-accession period. To do so, regression techniques and the instrumental variable method were adopted. This paper's main finding is the presence of a causal relationship between post-accession net migration rates and industry size in Poland, consistent with my theoretical analyses using the Specific Factor Model.

1. Introduction

Poland has historically been an emigrant, or net migration outflow, country. The fall of the Berlin Wall in 1989 coupled with an expansion of the European Union (EU) in 2004 magnified this trend by allowing for the free movement of Polish citizens across national borders. Post-accession, these conditions gave rise to one of the most significant population movements in contemporary European history. Official records for migration by the Polish Central Statistical Office (CSO), collected on the basis of permanent residence, reflect the drastic decrease in net migration (i.e., had more emigrants than immigrants), with the lowest number being -36,100 people in 2006, the lowest since the decommunization of Poland in 1989. While already significant, this number is often said to be an underestimation of the true scale of the Polish migratory movement, as it does not take into account Polish citizens who have moved abroad for long periods of time. This group of temporary migrants is significant because many Poles living abroad often choose not to give up their permanent residence as they perceive it as unnecessary or even disadvantageous. Based on estimations, the number of Polish emigrants rose by almost 1.5 million

people from 2004 to 2007, amounting to 6% of the total population.¹ In the field of international economics, many models (the Specific Factor model as an example) have been proposed to predict the impacts of migration. The case of Poland, with fluctuations in migration trends in the past two decades, lends itself to be an interesting real-world case study and prompts the question: how are population movements impacting the sending country's economy?

Ever since Poland's accession into the EU in 2004, many have attempted to answer this question. However, analysis has often been constrained to effects on the labor market (i.e., unemployment and employment, wages, and labor market efficiency). While the labor market may be most directly impacted by migration, industries are likely to be impacted as well. This paper investigates the effects of migration on two types of industries: labor-intensive industries and capital-intensive industries. Main findings suggest that there is a causal relationship between net migration rates and the growth of labor- and capital-intensive industries. Specifically, the greater the net migration, the bigger the industries in terms of gross value added to the Polish economy. This is significant in two ways: one, it explores the potential effects of migration outside the labor market, which is what the literature, in large part, has been confined to; second, it adds to the discussion on the effects of "brain overflow" to the Polish economy by providing quantitative data that shows the extent to which labor movements affect Polish industries (Kaczmarczyk and Okólski 2008, 621). Furthermore, it is also found that capital-intensive industries will be more heavily impacted by a loss in labor due to migration outflow of the working age population.

This paper is structured in the following order. The first section is an introduction of Polish post-accession migration. The second section will be comprised of an overview regarding the statistical backdrop of post-accession migration. The third will be a literature review of the existing research on the short-, medium- and long-term effects of migration to the Polish labor market. The fourth will be a discussion on the relationship between migration and the development of Polish industries from both theoretical and empirical perspectives. The fifth section will conclude on the findings of the paper and potential implications.

2. Statistics Regarding Post-Accession Migration

According to the literature, international migration between EU countries is extremely difficult to assess. This is especially true when it comes to temporary migrants who make up a significant portion of Polish migration outflow in comparison to other Central European countries (White, Grabowska, Kaczmarczyk, and Slany, 2018). To work around the difficulties surrounding the extent of migration, many researchers and government agencies like the CSO use available data to estimate the true scale and structure of Polish migration. As a result, the CSO is often described as having the most reliable source of data on post-accession migration in Poland (Kaczmarczyk 2012, 174). This paper will obtain most its data from the CSO. Not every piece of temporary emigration information can be estimated from available

¹ Own calculations based on: Kaczmarczyk, Pawel, "Labour market impacts of post-accession migration from Poland," (2012): 173-194; CSO--Central Statistical Office (2020), "Demographic situation in Poland up to 2019. International migration of population in 2000-2019," Updated March 12, 2020. <https://stat.gov.pl/en/topics/population/international-migration/>

data however, as a result, some data in this section will only consider migration for permanent residence. The use of such data could lead to a skewed perception of the scale of migration, and for this reason, I try to couple permanent residence migrant data with temporary migrant data. However, in cases where this is not possible, I believe migration based on permanent residence is still valuable when performing comparative analyses over periods of time.

As depicted in Table 1, migration outflow nets negative in the post-accession era. It should be noted that net migrant outflow might be overestimated given inconsistencies in the definition of temporary immigrants and emigrants. I do not believe, however, that the inconsistency will have any significant effect on the conclusion that net migration remains negative in Poland. Based on the table, temporary outflow is significantly larger than any other type of population movement. This is especially important because large temporary outflow is a distinctive feature of post-accession migration and what makes it different from other migration flows Poland experienced in the past. Thus, a more comprehensive review of temporary outflow is necessary to understanding post-accession migration movements.

Table 1. Net Migration in Poland (2010-2018)
(Author's own calculations based on data from the CSO, 2021)

Year	Inflow		Outflow		NET
	Permanent	Temporary	Permanent	Temporary	
2010	15200	155131	17400	2000000	-1847069
2011	15500	157059	19900	2060000	-1907341
2012	14600	217546	21200	2130000	-1919054
2013	12200	220311	32100	2196000	-1995589
2014	12300	222275	28100	2320000	-2113525
2015	.	218147	.	2397000	.
2016	13500	208302	12000	2515000	-2305198
2017	13300	209353	11900	2540000	-2329247
2018	15500	214083	11800	2455000	-2237217

To start, an overview on the magnitude of post-accession temporary outflow is necessary. By the end of 2006, two years after the accession, 1,950,000 Polish citizens were estimated to be temporary emigrants, i.e., 5.1% of Poland's then population. A year later, the number of temporary emigrants grew by 16.41% to 2,270,000 citizens.² Post-accession migration was nowhere near linear; it varied significantly based on the economic backdrop of the year, and is best demonstrated by following migration trends over the post-accession period.

² Ibid.

Table 2 below depicts the fluctuations in temporary emigration from 2000 to 2019. Starting from Poland's accession into the EU from 2004 to 2007, the number of temporary emigrants depict a steady upwards trend. From 2008 to 2009, however, a drop in temporary emigration can be observed. This slowdown could be attributed to the 2008 global recession, given that high unemployment rates throughout Europe decreased the incentive for migration. As the economy began to recover in 2010, Polish temporary emigration reverted to a heavy outflow. From 2017 to 2019, a decrease in temporary emigrants can once again be observed. This time, however, there were no indications of an economic downturn. Thus, decreased temporary emigration in times of economic prosperity demonstrates a possible transition of the Polish economy. Regardless of fluctuations, however, net migration rates remain negative.

Table 2. *Number of Temporary Emigrants in Thousands (From the CSO, 2020, and Kaczmarczyk, 2012)*

	Number of temporary emigrants in thousands															
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Overall	1000	1450	1950	2270	2210	1870	2000	2060	2130	2196	2320	2397	2515	2540	2455	2415
Within Europe	770	1200	—	—	—	—	1685	1754	1816	1891	2013	2098	2214	2241	2155	2134
Within the EU27	750	1170	1550	1860	1820	1570	1685	1670	1720	1789	1901	1983	2096	2121	2031	2008
Outside the EU27	20	30	—	—	—	—	78	84	96	102	112	115	118	120	124	126

In all four phases, distinctive features of post-accession migration can be drawn to differentiate it from past Polish migratory movements. One such feature is the age of post-accessions migrants. Based on data from the labor force survey interpreted by Kaczmarczyk (2012, 178), the median age of pre-accession migrants (30 years old) was higher than that of post-accession migrants (28 years old). The young age of temporary migrants indicates a work-based motive to emigrate. This is made apparent in Figure 1, where the temporary emigrant population is disproportionately comprised of working-aged people.

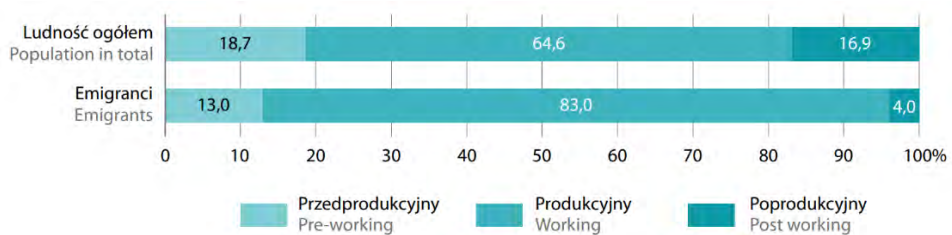


Figure 1. *Population in Total and Emigrants Staying Abroad for More Than 3 Months by Economic Age Group in 2011. (From the CSO, 2020)*

Before 2004, the unemployment rate in Poland was high due to an excess of labor (Kaczmarczyk 2012, 180). From 2001 to 2004, unemployment rates fluctuated between 16% to 19.7% (CSO, 2021). On balance, wages were also lower in Poland. In 2013, the average annual salary of a Polish citizen with no children was 7429 euros, whereas Germans and French earned three times more (Kolodziejczyk 2016, 23).

Thus, it can be reasonably inferred that many migrants moved abroad to seek more and better employment opportunities.

Another distinctive feature of post-accession migration is where, in relation to Poland at large, migrants come from. Historically, Polish emigrants came from a wide variety of geographical areas. However, recent migration patterns show a disproportionate increase in the number of migrants from economically underdeveloped regions with large contributions to agriculture and natural resource sectors (Kaczmarczyk, Mioduszevska, and Żylicz 2009; Kaczmarczyk and Okólski, 2008). It is, more specifically, educated citizens from these regions that have sought migration. Based on Table 3, compared to the pre-accession period where only 14.7% had university degrees, almost 20% of post-accession migrants had university degrees (Fihel and Kaczmarczyk, 2016). Furthermore, significantly more migrants came from urban areas than rural ones (CSO 2020, 85).

Table 3. Education Structure of Polish Pre- and Post- Accession Migrants by Gender (From Fihel and Kaczmarczyk, 2016)

Level of education	Percentage					
	Pre-accession			Post-accession		
	Total	Men	Women	Total	Men	Women
University degree*	14.7	12.0	18.3	19.8	15.6	27.0
Secondary	14.0	7.1	23.1	14.2	8.8	23.8
Secondary vocational	26.1	26.0	26.3	28.1	29.8	25.1
Vocational	34.8	45.4	20.9	30.9	39.2	16.2
Primary school	9.9	9.3	10.9	7.0	6.6	7.8
Unfinished primary school	0.4	0.2	0.5	0.0	0.0	0.0
Total	100.0	100.0	100.0	100.0	100.0	100.0

* Including bachelor's, master's, and Ph.D. degrees

In addition to where migrants came from, their country of destination also changed in the post-accession period. Based on the CSO (2020), in 2000, the number of permanent residence emigrants moving to Germany surpassed that of the second most popular country, the United States, by almost eight times. Nineteen years later, while Germany remained the most popular destination, the United Kingdom (UK) placed second and had 60% the number of permanent Polish emigrants that Germany had. Ireland also became a prominent destination and had the fastest emigrant growth rate (Kaczmarczyk and Okólski 2008, 603). The popularity of the UK and Ireland is further proven by Table 4, which depicts Polish migrant inflows in foreign countries from 2004 to 2014. In the post-accession period, the United Kingdom experienced a migration inflow of 328%, whereas Germany only experienced a percent change in Polish migration of 46.21%. Meanwhile, Ireland experienced a growth of 667% in ten years. Thus, it can be concluded that most Polish temporary emigrants in the post-accession period sought labor opportunities in the UK and Ireland instead of Germany, the traditional destination.

Table 4. *Most Popular Destinations for Polish Migrants in the EU*
(Author's own calculations based on data from Kolodziejczyk, 2016)

Country	Number of migrants at the end of 2004	Number of migrants at the end of 2014	% Change
United Kingdom	150	642	328.00%
Germany	383	560	46.21%
Ireland	15	115	666.67%
The Netherlands	23	113	391.30%
Italy	59	96	62.71%
France	30	63	110.00%
Norway	.	71	.
Belgium	3	49	1533.33%
Sweden	11	40	263.64%

(.) indicates no data, numbers in thousands.

Reasons for the influx of emigrants to the UK and Ireland are as follows: one, the UK and Ireland were countries where many migrants felt they could earn additional income; second, potential migrants received recommendations from those who already worked there; third, many respondents also stated choosing the UK or Ireland to learn the language; lastly, the UK and Ireland, being a part of the EU (at the time) were more easily accessible to Polish migrants than alternatives like the US.

3. Literature Review

The effects of migration on the labor market in post-accession Poland have been studied comprehensively, and this section aims to provide context by summarizing generally agreed upon conclusions in the literature as aforementioned. Described as a “natural experiment”, post-accession migratory movements have been of great interest in the literature (Kaczmarczyk, Aldaz-Carroll, and Holda, 2020). To more comprehensively analyze the effects of net outflow on the economy, studies have often approached the question by analyzing migration’s short-, medium-, and long-term effects. Given the literature’s strong focus on the effects within the labor market, the migration theory, which helps link certain effects of migration with specific timeframes, is a framework that has previously been used (Kaczmarczyk 2012, 179-180). Under this theory, short-term effects are associated with supply and demand, therefore referring to changes in employment and unemployment. Medium-term effects are related to the structural features of the migration (e.g., potential brain drain) that take place, as well as a set of market adjustments, that could, among other things, lead to changes in wages. The long-term effects are associated with the possibility of an additional set of market adjustments, including changes in the structure of the economy (e.g., capital/labor ratio), social and occupational mobility of domestic workers, and immigration of labor. The section will be split into three parts: the first on short-term impacts, the second on medium-term impacts, and the third on long-term impacts.

Based on the migration theory, one of the most important and relevant short-term indicators in the case of migration is unemployment and employment (i.e., the number of jobs available). Many researchers acknowledge a decline in unemployment in the first phase of post-accession migration (Kolodziejczyk, 2016; White, Grabowska, Kaczmarczyk, and Slany 2018; Kaczmarczyk 2012). Kaczmarczyk argues, however, that there may be little causality between post-accession migration and

unemployment throughout the post-accession period. He offers three reasons for this: first, based on a Pearson correlation coefficient of -0.97 , unemployment in the first phase of post-accession migration is strongly correlated with increases in employment; second, unemployment and employment trends in the labor market continue throughout the second phase of post-accession migration (2007-2008), where migration rates face a decline; third, the increase in emigrants of approximately 300,000 is significantly less than the fall in unemployment by 2 million (Kaczmarczyk 2012, 181). Lo Turco and Parketa's (2008) findings that new EU states experienced increased employment (when their EU15 trade partners also underwent an increase in employment) support Kaczmarczyk's claim that unemployment and employment was shaped by the economic situation at large. There is no proof that post-accession migration is not at all correlated with employment, just as there is no proof that post-accession migration was the main driver of changes in employment and unemployment. Therefore, the consensus in literature is that in the short-term, post-accession migration (phases 1, 2, and 3) had a moderate impact on the Polish economy.

In the medium term, structural features of the migration are taken into consideration. Brain drain is often presented as a possible concern given that post-accession migration pre-2017 is characterized by a large outflow of educated young citizens. Kaczmarczyk, however, points out that while employers faced a growth in labor shortages in the first phase of post-accession migration (2004-2007), they ceased to be an issue by 2008. Furthermore, shortages of labor were mostly discernable in the construction and manufacturing sector, with 35% and 15% of firms affected respectively. While these workers were qualified, they would not necessarily be described as highly skilled (White, Grabowska, Kaczmarczyk, and Slany 2018, 100). Thus, one might argue that Poland was not experiencing "brain drain" but a "brain overflow" (White, Grabowska, Kaczmarczyk, and Slany, 2018; Kaczmarczyk, Mioduszewska and Żylicz 2009; Kaczmarczyk and Okólski 2008; Kahanec and Zimmermann 2008). The literature agrees that brain overflow would not drastically impact the Polish economy as brain drain would, but it is extremely difficult to prove that brain overflow will have no impact at all. Thus, in the medium term, the literature generally agrees on a moderate impact.

Post-accession migration in Poland is, comparative to history, a very recent phenomenon. As such, at the time of most significant outflow (phase 1) where the greatest number of studies were written on migration's impact to Poland's economy, one could only speculate on long-term impacts. The most frequently mentioned theory is the modernization of the Polish economy through massive outflow—proposed by Kaczmarczyk and Okólski and coined the "crowding-out" hypothesis (Kaczmarczyk and Okólski, 2008). In an analysis of previous patterns of European economic development, Blanchard, Rudiger, Krugman and Layard (1992) propose a relationship between the massive outflow of Western and Northern Europeans overseas to economic development in these regions. Based on Layard's idea, Kaczmarczyk and Okólski argue that for a variety of reasons, Eastern Europe was unable to attain the crowding-out effect in the nineteenth and twentieth centuries. As a result, countries faced issues such as having greater population sizes than the economy needed. The accession into the EU, however, provided an opportunity for citizens from economically underdeveloped regions to emigrate to more economically developed countries with little-to-no barriers. This resulted in significant population outflow and would be predicted to create an incentive for improving the efficiency of domestic labor markets. Thus, while migration flows in the post-accession period only had moderate impacts in the short and medium term, the crowding-out effect

could potentially lead to changes in the structure of the labor market, allowing Poland to develop in the long term (Kaczmarczyk and Okólski 2008, 614-621; Kaczmarczyk 2012, 186-189).

4. Impacts of Migration on Polish Industries

The most direct impacts of population movements on the economy can be found in the labor market. However, there is also reason to believe that migration is also correlated to the performance of industries in the medium term (i.e., after a set of market adjustments, most specifically wages, have taken place). This is because industries are comprised of firms who need people to act as consumers, driving demand, and suppliers of labor, allowing for greater production. There is, however, little research done on the impacts of post-accession migration to Polish industry growth. This section aims to shed light on this topic by analyzing the impact that post-accession migration has on labor-intensive and capital-intensive industries.

Theoretical Analysis

This analysis can apply the Specific Factor Model, a variant of the Ricardian model where factors of production are assumed to be specific to a certain industry. First, the Specific Factor Model assumes that capital is specific to a particular industry and thus immobile, making it applicable to medium-term analyses. Meanwhile, labor is assumed to be freely mobile, an important assumption in an analysis where the aim is to measure the effect migration has on the economy. From the assumption that some factors of production are more mobile across certain industries than others, the Specific Factor Model's design makes it the most applicable in the real world and is therefore adopted in this analysis (Melitz, Obstfeld and Krugman, 2012).

To set up the Specific Factor model, Polish industries are grouped in two overarching categories: labor-intensive industries, O_L , and capital-intensive industries, O_K . In this analysis, O_L and O_K will be treated as two collective industries.

Figure 2 depicts the Specific Model for Poland in pre-migration conditions. Point L indicates the amount of labor employed in equilibrium while Point K indicates equilibrium wage.

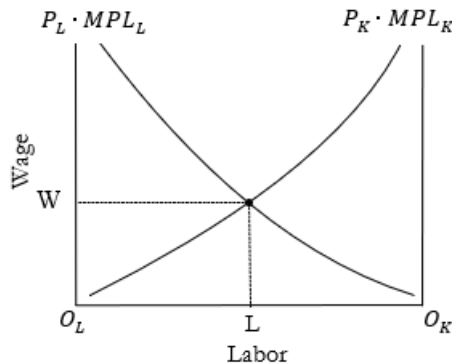


Figure 2. *Specific Factor Model of Poland Pre-Migration*

Figure 3 below shows the Specific Factor Model in Poland accounting for population movements. Referring to Table 1, Poland has maintained a negative net migration outflow from 2010 to 2018. While net migration outflow has been trending positive (i.e., a smaller negative net migration) from 2016 to 2018, it is unlikely that Poland has attained positive net migration at the time this paper is written. As previously mentioned, many temporary emigrants (who contribute greatly to the negative net migration) are young, educated adults who would otherwise have contributed to the Polish workforce. Thus, a leftward shift of the graph is shown with the blue line to reflect negative net migration. Points L' and W' illustrate new points of equilibrium labor employed and wages respectively.

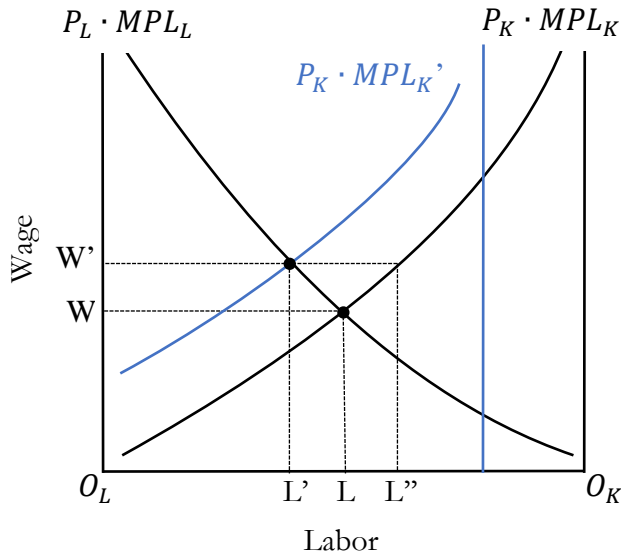


Figure 3. *Specific Factor Model in Poland After Net Outflow*

From Figure 3, two observations can be made: first, labor-intensive industries experience a loss in the amount of labor employed as L shifts to L' ; second, capital-intensive industries also experience a loss in labor as a result of L shifting to L'' . From these two observations, net migration outflow causing both capital-intensive and labor-intensive industries to decrease in size in the medium-term can be concluded given the assumption that labor is a prerequisite for industry growth.

Empirical Analysis

Models, being a gross simplification of the real world, are beneficial for identifying relationships but unreliable as the only basis for drawing real world conclusions. Thus, to test the conclusion in the previous section, empirical analyses on available data are necessary. In this case, regression analyses were used to identify the correlation of net migration outflow on the gross value added of labor-intensive industries (O_L) and capital-intensive industries (O_K). The regression analyses spanned from 1995 to 2020 and contained 26 observations.

Three industry groups were chosen for both O_L and O_K . For O_L , the industry groups selected were the agriculture/forestry/fishing industry group, mining/quarrying industry group, and construction industry group. For O_K , the manufacturing industry group, electricity/steam/gas/air conditioning supply industry group, and transportation industry group were used. The gross value added (GVA) of the respective industry groups was then added up to form a collective GVA for O_L and O_K . GVA was chosen as the metric because it accounts for the dollar value of goods and services provided by the industry without taking taxes and subsidies into consideration, which is a more accurate indicator of the size of O_L and O_K compared to other indicators like GDP (Kenton, 2021).

The econometrics models are as follows:

$$GVA_{O_L} = \beta_0 + \beta_1 \text{net migration rate} + \beta_2 \text{population} + \beta_3 \text{life expectancy} + u$$

$$GVA_{O_K} = \beta_0 + \beta_1 \text{net migration rate} + \beta_2 \text{population} + \beta_3 \text{life expectancy} + u$$

Using the models above, the regression analyses yield the following results:

Table 5. Regression Statistics for GVA_{O_L}
(Based on author's own calculations using data from the CSO, 2021 and macro trends)

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-4934909.2	1208167.103	-4.084624686	0.00049066	-7440494.39	-2429324	-7440494.4	-2429324
Net Migration Rate	35972.2522	14608.69672	2.462386133	0.02210569	5675.669533	66268.8349	5675.66953	66268.8349
Population	0.07425125	0.032812913	2.262866888	0.03385704	0.006201438	0.14230107	0.00620144	0.14230107
Life Expectancy	30081.0838	2588.966263	11.61895551	7.3692E-11	24711.89642	35450.2712	24711.8964	35450.2712

Note: R squared for this data set is 0.91509007; Significance F is 6.14424E-12.

Table 6. Regression Statistics for GVA_{O_K}
(Based on author's own calculations using data from the CSO, 2021 and macro trends)

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-11193389	2609356.352	-4.289712739	0.00029733	-16604863	-5781915.3	-16604863	-5781915.3
Net Migration Rate	154626.293	31551.34375	4.900783127	6.701E-05	89192.811	220059.775	89192.811	220059.775
Population	0.12789265	0.070868163	1.804655958	0.08483658	-0.0190789	0.27486423	-0.0190789	0.27486423
Life Expectancy	89383.7498	5591.557281	15.98548406	1.3562E-13	77787.5697	100979.93	77787.5697	100979.93

Note: R squared for this data set is 0.9438254; Significance F is 4.0878E-13.

P-values of less than 0.05 are present in both Table 5 and 6, meaning that the null hypothesis most likely does not hold true. In this case, it can be concluded that the net migration rate is correlated with the size of labor- and capital-intensive industries.

The establishment of this correlation indicates that the coefficients can be further investigated. In doing so, two observations can be made:

- For labor-intensive industries, an increase of 1 in the net migration rate leads to an increase of roughly 35,972 2020 zloty in GVA.
- For capital-intensive industries, an increase of 1 in the net migration rate leads to an increase of roughly 154,626 2020 zloty in GVA.

From these two observations, it can be said that in the real world, both labor- and capital-intensive industries become smaller when there is an influx of labor outflow. In particular, capital-intensive industries suffer a greater loss than labor-intensive industries in GVA. Because shortages of labor were most present in the construction and manufacturing sector of the Polish economy (with 35% and 15% of firms affected respectively, as stated previously in Part 2 of this paper), it can be proven empirically (as it has been done theoretically in the first part of the analysis) that labor shortages are at least a partial mechanism for the correlation between net migration rates and the size of capital- and labor-intensive industries.

Tables 5 and 6 affirm a positive correlation between net migration and the growth of both labor- and capital-intensive industries. Proving a correlation, however, does not imply causation. Thus, this paper adopts the instrumental variable analysis to infer potential causal relationships between net migration rate and the growth of labor- and capital-intensive industries. In the analysis, x will stand for net migration rate, y for GVA, and z for the instrumental variable: median gross weekly earnings in the UK. Figure 4 depicts this relationship.

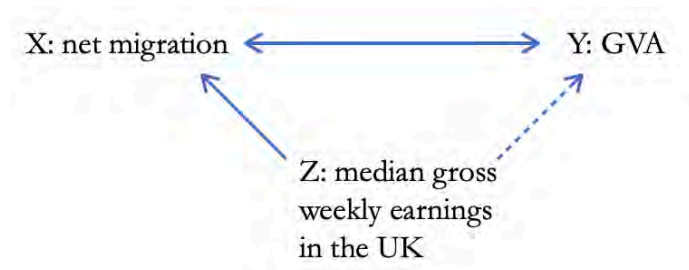


Figure 4. Relationship between variables X , Y , and Z

The instrumental variable chosen in this analysis are the median gross weekly earnings in the UK. This variable was chosen for the following reasons. First, the instrumental variable analysis method assumes that there is no correlation between the instrument (variable Z) and the dependent variable (variable Y) except through the independent variable (variable X). Earnings in the UK are highly unlikely to be directly correlated with industry growth in Poland as there is a small likelihood that labor market conditions in one country directly affect the industry growth of another. Second, because Poland is experiencing net outflow (Table 1) where the majority are temporary migrants, weekly earnings seemed most indicative of wages migrant workers would earn as opposed to annual earnings. Third, the UK was chosen specifically because of its increasing popularity as a Polish migrant destination spot post-accession (Table 4). Therefore, median gross weekly earnings in the UK (in £) was chosen as the instrumental variable because it both fulfills a major assumption underlying the instrumental variable analysis method and is representative of wage conditions the majority of post-accession Polish migrants would experience.

The first step for both O_L and O_K in the instrumental variable analysis is the econometrics model as follows:

$$\text{net migration rate} = \beta_0 + \beta_1 \text{median gross weekly earnings in the UK} + \beta_2 \text{population (Poland)} + \beta_3 \text{life expectancy (Poland)} + u$$

Using the model above, the regression analysis yields:

Table 7. Regression Statistics for the Correlation of Variable Z on X
(Author's own calculations using data from the CSO, 2021 and the Office for National Statistics, 2020)

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	34.3116309	14.61235186	2.348125147	0.02926527	3.830799	64.7924627	3.830799	64.7924627
Median Gross Weekly Earnings (UK) in £	0.00781572	0.002733921	2.838793993	0.0097054	0.00211286	0.01351858	0.00211286	0.01351858
Population (Poland)	-2.217E-07	4.62705E-07	-0.479108232	0.63705842	-1.187E-06	7.435E-07	-1.187E-06	7.435E-07
Life Expectancy (Poland)	-0.3925462	0.114656427	-3.423673967	0.00269032	-0.6317153	-0.1533771	-0.6317153	-0.1533771

Note: Significance F is 0.00034689.

Given a P-value of 0.0097 (which is less than 0.05) for median gross weekly earnings in the UK, it can be concluded that the instrument is correlated to the independent variable. Using the coefficients of the intercept and median gross weekly earnings, an equation can be used to find *net migration'*, a prediction of the probability of net migration based only on median gross weekly earnings in the UK. The equation is as follows:

$$\text{Net Migration}' = \beta_0 + \beta_1 \cdot \text{median gross weekly earnings in the UK}$$

Using the data points obtained from the equation above, a second econometrics model can be used where:

$$GVA_{OL} = \beta_0 + \beta_1 \text{net migration rates}' + \beta_2 \text{population (Poland)} + \beta_3 \text{life expectancy (Poland)} + u$$

or

$$GVA_{OK} = \beta_0 + \beta_1 \text{net migration rates}' + \beta_2 \text{population (Poland)} + \beta_3 \text{life expectancy (Poland)} + u$$

Using the models above, the regressions yield:

Table 8. Regression Statistics for the Correlation of Variable X' on Y for GVA_{OL}
(Author's own calculations using data from the CSO, 2021, the Office for National Statistics, 2020, and Table 7)

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-8017740.3	737169.1566	-10.87639137	7.5656E-10	-9555448.2	-6480032.3	-9555448.2	-6480032.3
Net Migration Rate'	129576.875	13579.49149	9.542100654	6.9171E-09	101250.552	157903.197	101250.552	157903.197
Population (Poland)	0.11125158	0.017962664	6.193490163	4.7482E-06	0.07378212	0.14872104	0.07378212	0.14872104
Life Expectancy (Poland)	-12144.287	4451.073538	-2.72839516	0.01294667	-21429.064	-2859.5108	-21429.064	-2859.5108

Note: R squared for this data set is 0.97611833; Significance F is 2.209E-16.

Table 9. Regression Statistics for the Correlation of Variable X' on Y for GVA_{OK}

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	360393.302	3796935.116	0.094916898	0.92532539	-7559874.6	8280661.16	-7559874.6	8280661.16
Net Migration Rate'	0.15071827	0.065482492	2.301504768	0.03224668	0.01411418	0.28730235	0.01411418	0.28730235
Population (Poland)	-0.1563608	0.10384199	-1.505757146	0.14776074	-0.3729714	0.06024978	-0.3729714	0.06024978
Life Expectancy (Poland)	78948.6118	8231.138212	9.591457438	6.3502E-09	61778.7584	96118.4653	61778.7584	96118.4653

Note: R squared for this data set is 0.89950597; Significance F is 3.705E-10

As the P-value for *Net Migration Rate'* is significantly less than 0.05, it is said to be statistically significant. Based on the assumption that there is no theoretical correlation between Variable Z and Variable Y, it can be concluded that there is a causal pathway (or, at the very least, disproves correlation due to reverse causality) between net migration and GVA_{OL} and GVA_{OK} . Therefore, the empirical analysis supports the theoretical conclusion that net migration outflow causes labor- and capital-intensive industries to decrease in size in the short run.

5. Conclusion

High unemployment rates in Poland and the ability to move across EU nations with little-to-no regulation after 2004 prompted large outflows of migrants in Poland. Referred to broadly as the post-accession migration movement, it is characterized by a large number of temporary migrants (of which the number varied depending on the economic background at the time) who were young, educated, and come from less economically developed regions of Poland. There was also a shift in the country of destination. Pre-accession, Germany was the most popular country of destination, but post-accession, English-speaking European countries like England and Ireland gained popularity. Because of its magnitude, post-accession migration has been broadly studied – especially when it comes to its impact on the labor market. Applying the migration theory, it is generally agreed upon in the literature that short- and medium-term impacts are moderate whereas long-term impacts could include incentivizing a change in Poland's labor structure through the "crowding-out" effect.

The main finding of this paper is the establishment of causation between net migration and labor- and capital-intensive industries in Poland. I find that with increased population inflow and/or decreased outflow, both labor- and capital-intensive industries experience growth depicted by an increase GVA. The relationship was first studied theoretically using the Specific Factor Model, which yielded the conclusion of a positive correlation between net migration and industry size (i.e., net immigration causes growth in both the capital- and labor-intensive industry and vice versa). Empirical analyses of real-world data were then conducted to investigate the relationship between the net migration rate and industry size. The first two empirical analyses found evidence of a correlation between net migration rate and industry size. To establish causation, the instrumental variable analysis method was used. Given that *Net Migration Rate'* was found to be statistically significant for both GVA_{OL} and GVA_{OK} , the empirical analysis corroborated the Specific Factor Model, and a causal relationship between net migration rate and industry size can be established.

The causation between net migration and industries in Poland is promising, as it is indicative of the potential for future research to be done in areas of the economy that are indirectly affected by migration. It reveals the complexities of the real-world economy and how one change, such as a change in population, can have effects on multiple aspects of a country's economy.

Bibliography

- Blanchard, Oliver, Rudiger Dornbusch, Paul Krugman, and Richard Layard. *East-West migration: the alternatives*. Mit Press, 1992. <http://mitpress.mit.edu/main/home/default.asp>
- CSO--Central Statistical Office. "Demographic situation in Poland up to 2019. International migration of population in 2000-2019. Updated March 12, 2020. <https://stat.gov.pl/en/topics/population/international-migration/demographic-situation-in-poland-up-to-2019-international-migration-of-population-in-20002019,6,1.html>
- CSO--Central Statistical Office. "Annual macroeconomic indicators." Updated September 3, 2021. <https://stat.gov.pl/en/poland-macroeconomic-indicators/>
- Fihel, Agnieszka, and Paweł Kaczmarczyk. "Migration: a threat or a chance? Recent migration of Poles and its impact on the Polish labour market." In *Polish migration to the UK in the 'new' european union*, pp. 23-48. Routledge, 2016.
- Kaczmarczyk, Paweł. "Labour market impacts of post-accession migration from Poland." *Free Movement of Workers and Labor Market Adjustments: Recent Experiences from OECD Countries and the European Union*, OECD Publishing (2012): 173-194. <https://doi.org/10.1787/9789264177185-10-en>
- Kaczmarczyk, Paweł, and Marek Okólski. "Demographic and labour-market impacts of migration on Poland." *Oxford Review of Economic Policy* 24, no. 3 (2008): 599-624. <https://doi.org/10.1093/oxrep/grn029>
- Kaczmarczyk, Paweł, Enrique Aldaz-Carroll, and Paulina Hołda. "Migration and Socio-economic Transition: Lessons from the Polish Post-EU Accession Experience." *East European Politics and Societies* 34, no. 4 (2020): 910-936.
- Kaczmarczyk, Paweł, Marta Mioduszevska, and Anna Żylicz. "Impact of the post-accession migration on the polish labor market." In *EU labor markets after post-enlargement migration*, pp. 219-253. Springer, Berlin, Heidelberg, 2009.
- Kahanec, Martin, and Klaus F. Zimmermann. "International migration, ethnicity and economic inequality." *Oxford: Oxford University Press*. 455–90. (2008).
- Kenton, Will. "Gross Value Added (GVA)." Investopedia. Updated June 29th, 2021. Accessed August 31st, 2021. <https://www.investopedia.com/terms/g/gross-value-added.asp>
- Kolodziejczyk, Katarzyna. "Poland in the European Union. Ten years of membership." *Revista UNISCI* 40 (2016): 9-26.
- Lo Turco, Alessia, and Aleksandra Parteka. "EU Enlargement, Economic Interdependence and the Labor Markets in 'Old' and 'New' Member States." *Available at SSRN 1337078* (2009).
- Macrotrends. "Poland Net Migration Rate 1950-2021." Accessed August 30st, 2021. <https://www.macrotrends.net/countries/POL/poland/net-migration>
- Melitz, Marc, Obstfeld, Maurice, and Krugman, Paul. *International Trade: Theory and Policy*. Saylor Academy, 2012. E-book. https://saylordotorg.github.io/text_international-trade-theory-and-policy/s08-15-the-specific-factor-model-over.html

- ONS--Office for National Statistics. "Annual Survey of Hours and Earnings time series of selected estimates". Released November 3, 2020. Accessed September 4, 2020. <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/ashe1997to2015selectedestimates>
- White, Anne, Izabela Grabowska, Paweł Kaczmarczyk, and Krystyna Slany. *The impact of migration on Poland: EU mobility and social change*. UCL Press, 2018. DOI:10.2307/j.ctv550d7m

