QuaesitUM

The Undergraduate Research Journal of The University of Memphis

Spring 2025

http://www.memphis.edu/quaesitum/

QuaesitUM is an annual publication that provides an academic forum where University of Memphis undergraduate students can showcase research from all disciplines.

ISSN 2375-5423
Printed in Canada

Editorial Board

Chaos is the teacher that forces us to adapt and learn. In the face of chaos, find your own truth and stand firm.

–Unknown

To Our Readers

Each year, as we prepare *QuaesitUM* for press, I find myself looking reflectively on the University of Memphis and the climate in which research that is represented here was conducted.

The 2024-2025 academic year has brought tremendous upheaval, as social structures have been dismantled and support for research has waned on a national scale. The very foundations of education have been challenged, and we have been forced to cling to what we once took for granted. Efforts to find out more about the world – what it is to be human, what will make humanity better, what will enhance our lives – have come under fire from some quarters. It is in this uncertain environment, however, that research is most critical. It is the journey through chaos and uncertainty – making sense of the madness – that leads to our greatest discoveries.

We chose the name *QuaesitUM*, meaning to seek or to know, twelve years ago because it captured our vision of what a research journal, and indeed a university, should be. Seeking is the journey through chaos – an attempt to make sense of what is currently unknown. And knowing is our resting place once we have made sense of the disorder.

The papers included here reflect important steps on that journey through turmoil toward understanding and knowledge. They provide important pieces in the chaotic puzzle of the greater world. Despite an environment where some might try to stifle their investigations, these researchers have continued to battle the upheaval and persevere – to ask the questions that lead to discovery.

As has been the case for the past twelve years, I have once again been privileged to have the partnership of Dr. Melinda Jones, my co-editor and the Director of the Helen Hardin Honors College. Each year, I am more grateful to have her as a collaborator. Heri Yusup, PhD candidate in Applied Linguistics, has also been a vital part of our team for the past two

years. His commitment to ensuring the quality and rigor of the journal has been pivotal in creating the best possible representation of our students' work. This year we also added two Mellon Interns to our team as student editors: Donice Jenkins and Timothy Presley. Their insights and perspectives made them valuable additions to our team.

Professor of Graphic Design Gary Golightly continues to inspire with his innovative and eye-catching cover designs. For the past twelve years, his artwork has captured the spirit of the journal and the papers therein. Thanks are also due to administrators who have offered support – Dr. Abby Parrill-Baker, Dr. Thomas Nenon, Dr. Karen Weddle-West, and Dr. Jasbir Dhaliwal, and the faculty reviewers who gave freely of their time and expertise in evaluating the authors' work.

The journal would not be possible, moreover, without the efforts of the faculty sponsors. Dr. Rebecca Howard, Dr. Nicole Detraz, Dr. Lindsey Feldman, Dr. Max Garzon, Dr. Thomas Hagen, Dr. Carl Herickhoff, and Dr. Chrysanthe Preza inspired and guided their students in their research, helping them journey into chaos and emerge with greater knowledge.

As always, however, the most important contributors are the student authors themselves. David Adaway, Addison Cucchiaro, Gerrell Dabbs, Hema Dalavayi, Farraday Johnson, Jackson Lyles, Karma Rakow, and Jacob Stewart took on the daunting task of asking difficult questions and seeking answers despite uncertainty. What they have produced leads us toward greater understanding of the world around us.

With gratitude to all,

Dr. Sage L. Graham
QuaesitUM Editor-in-Chief

Table of Contents

*This paper received a *QuaesitUM* outstanding paper award.

Publication and Review Process

Addison Cucchiaro is currently pursuing a bachelor's degree in Art History with a minor in Italian Studies at the University of Memphis. She is on track to graduate in May of 2025 and will be attending the graduate program in Art History at American University in Washington, D.C. this fall. Addison's long-term goal is to earn a PhD in Art History, with a focus on the Italian Renaissance and Baroque periods. Addison's decision to submit her work to *QuaesitUM* was motivated by a desire to share her academic findings with a wider audience. Her research, particularly in the fields of gender and queer studies in art history, highlights the intersections of art and identity, exploring the ways in which gender and sexuality have been represented in art history. Her research aims to bring forward these underrepresented perspectives and challenge traditional narratives.

Addison's paper received a *QuaesitUM* outstanding paper award.

Addison Cucchiaro

Sappho and Erinna in a Garden at Mytilene
and the Art History of Classical Lesbian Love

Faculty Sponsor
Dr. Rebecca Howard

Abstract

This research examines the works of Victorian artist Simeon Solomon, focusing on his representations of the ancient poetess *Sappho*—the earliest documented lesbian figure in Greco-Roman history. The research explores Solomon's personal life and analyzes his artwork through an iconographic lens, incorporating feminist and queer theory. His depictions of *Sappho* are also compared to other portrayals of classical lesbian figures, particularly those that are overly sexualized. By doing so, this research aims to open a conversation about the complex history of lesbian love in art history.

...and among all animals, not one female is attacked by lust for a female – Ovid, Metamorphoses

Introduction

The field of art history has undoubtedly undergone needed modification over the course of the last several decades. Though the view of the straight, white man continues to reign supreme, broader narratives have surfaced and earned respect. Studies in race, gender, and sexuality have proved their importance to the expansion of the field. Yet, there is still much to be said about the representation of certain minority groups. For instance, queer women have been exceptionally underrepresented throughout the history of art. Although we cannot go back in time and cure the world of both misogyny and homophobia, and therefore grant access to the artist to freely paint such themes as queerness between women, what we *can* do now is look back at the handful of artistic depictions of queer women and strengthen their importance to the understanding of lesbianism's history. A good place to begin analysis is Simeon Solomon's *Sappho and Erinna in a Garden at Mytilene*. By examining Solomon's personal life, looking at the piece through an iconographical lens, applying feminist and queer theories, and comparing it to other paintings that portray overly sexualized classical lesbian figures, we can begin to prioritize a conversation surrounding the complex history of ancient lesbian love as depicted in art.

Simeon Solomon and the Art of Victorian England

Simeon Solomon was born on October 9th, 1840, in London, England, to a large middle class Jewish family. He was the youngest of Michael and Kate Levy Solomon's eight children. His interest in the arts was evident from a young age, having been heavily influenced by his mother, an amateur miniatures artist, his brother Abraham, and his sister Rebecca, both of whom were successful working artists. In April of 1856, at the age of sixteen, he began to attend the Royal Academy of Art in London, where his main artistic influences showed to be from both Shakespearean and Judaic texts. Another obvious influence on his early works was art from the Pre-Raphaelite Brotherhood.[1] The Pre-Raphaelite Brotherhood was formed in 1848 by a group of young artists reacting against the stricter academic styles coming out of their training institution, the Royal Academy

[1] Roberto C. Ferrari, "Solomon's Life Before 1873," Simeon Solomon Research Archive, accessed April 19, 2024, https://www.simeonsolomon.com/

of Art, which had long dominated the art scene in London. The art of Raphael (born Raffaello Sanzio da Urbino) of the Italian High Renaissance was the primary influence for the students of the Academy, whose artistic approaches were concluded to be mundane and formulaic by the Brotherhood. The Pre-Raphaelites favored the simplicity of line, the use of bright coloring, a meticulous attention to detail, and the use of medieval subject matter as seen in art before the ascendancy of Raphael.[2] Although students of the Academy, young artists like Solomon were evidently very impacted by the rebellious styles of the Pre-Raphaelites. Solomon most likely met Dante Gabriel Rossetti, one of the founders of the Brotherhood (along with William Holman Hunt and John Everett Millais), around 1858. It was during this same year that his first Royal Academy work was exhibited, *Isaac Offered*, along with two drawings at the Ernest Gambart's Winter Exhibition.[3] He soon became well known amongst the Pre-Raphaelite circle and formed notable colleagueship with Rossetti, Everett Millais, Holman Hunt, and Edward Burne-Jones. Through either Rossetti or Burne-Jones, Solomon was introduced to poet and critic Algernon Charles Swinburne, whose written work notoriously favored classical themes.[4]

The classical world has long been one of the artist's greatest muses (perhaps second to the biblical world). Victorians sought to create art (whether that be painting, literature, or any other artistic expression) that would be celebrated for generations to come; they wanted to be remembered for their greatness, or else their art was utterly insignificant. So, they looked to the greatest period of art thus far: the Italian Renaissance. In Italy, during the Quattrocento (the fifteenth century), the resuscitated knowledge of classical antiquity, recovery of its lost culture, and the renewed dispersion of a formally dormant liberal spirit came to be what we know as the Renaissance.[5]The idea of Humanism was re-birthed, which revived the pagan views of the ancient Greeks and Romans, who maintained that "man was made, not only to toil and suffer, but to enjoy."[6] Artists during the Victorian age looked back to the Renaissance for inspiration, and therefore the ancient world. With the rise of first-hand observation of antique artworks and widespread reproductions, the classicism of the Renaissance

---

[2] Rebecca Jeffrey Easby, "A beginner's guide to the Pre-Raphaelites," Smarthistory, last modified August 9, 2015, https://smarthistory.org/a-beginners-guide-to-the-pre-raphaelites/

[3] Ferrari, "Solomon's Life Before 1873."

[4] Ibid

[5] A.W. Ward, G.W Prothero and Stanley Leathes, "The Classical Renaissance," The Cambridge Modern History 1, Cambridge University Press (1907), 532.

[6] Ibid, 533

had a revision.[7] Neoclassicism was the new wave of the Renaissance's classicism; another renewal of the styles and spirit of the classical world.

Soon after befriending Swinburne, Solomon began to incorporate Greco-Roman themes into his art, as well as homoerotic. Both Solomon and Swinburne had what was deemed unconventional sexualities for the time period (meaning they were outwardly homosexual), and they expressed their resulting internal turmoil artistically. Though there is no evidence to confirm any sort of romantic relationship between the two men, the influence that Swinburne had on Solomon's art suggests a close friendship between the two, or at the very least a profound professional interest in one another.

Solomon's new classical style was tied to the emerging Aestheticism movement in the arts, which was promoted by Swinburne. The Aesthetic Movement, which can be seen as both an extension of the Pre-Raphaelite Brotherhood and a renewal of some of their concepts, was characterized by the expression 'art for art's sake.' Artists following this movement prioritized the importance of beauty in their artworks, forgoing the idea that art must have a legible narrative or moral message.[8] Art was simply, for lack of better words, art; something to be looked upon and admired. Which is why, when face to face with artworks of this era, one would meet many beautiful women looking back at them. The male gaze, which is defined as the "sexually objectifying attitude that a representation takes toward its feminine subject matter, presenting her as a primarily passive object for heterosexual-male erotic gratification," was at a crux during the Aesthetic Movement, as the epitome of natural beauty was regarded as an idealized, youthful woman.[9] Within his early biblically-themed works, it is evident that Solomon began to explore his sexuality artistically. There are several sketches showing the figures of David and Jonathan embracing and kissing. Many of these works eroticize the male body, counteracting the works coming out of the Pre-Raphaelites that eroticized women. Solomon was considered unique among those following Aestheticism, due to his use of "effete" and "languorous" male figures.[10] Unlike his heterosexual counterparts, Solomon's art did not overly sexualize the female body, making his

---

[7] Ibid

[8] Rebecca Jeffrey Easby, "The Aesthetic Movement," Smarthistory, last modified June 3, 2016, https:// smarthistory.org/the-aesthetic-movement/

[9] A.W Eaton, "Feminist Philosophy of Art," Philosophy Compass 3, no. 5 (September 2008): 878.

[10] Ferrari, "Solomon's Life Before 1873."

portfolio stand out against others from the time. He strayed from the classical subjects that his contemporaries favored, like Venus. Instead, he often used androgynous male deities as subjects, such as Bacchus and Cupid.[11] It may, then, initially seem odd that he completed a piece depicting two women engaged in a seemingly homosexual act. Throughout the 1860s, Swinburne wrote and published several poems regarding Sappho. I propose that this influence, paired with the time period's artistic emphasis on ideally beautiful female subjects, prompted Solomon to use Sappho as an outlet for his own expression. The Sappho in which both Solomon and Swinburne depict in their artistic narratives became the starting point of "the emergence of the modern image of Sappho as a Lesbian."[12] In search of a Muse, an outlet, both Swinburne and Solomon found Sappho.

Sappho of Lesbos

The modern terms "Sapphic" and "Lesbian" both describe homosexual relationships between women, and they are both in clear reference to the existence of Sappho and her documented romances with other women. "Sapphic" is a clear alteration of the woman's name, and "Lesbian" derives from the fact that Sappho was born on the island of Lesbos, where the inhabitants were often called "Lesbians". The modern definition was popularized in the late nineteenth century as the label used by sexologists to describe "love between women as an abnormal condition, a pathology."

Although today's culture sees Sappho as an icon for female homosexuals, before the nineteenth century this was not so. In a classical Athenian comedic play, Sappho was portrayed as a promiscuous heterosexual woman who had tumultuous relationships with several ancient men, some of whom were not even alive during her lifetime. It is assumed that, within the two centuries between her death and the start of her "literary after-life" in Athens, "there was an attempt to make up for long lost knowledge about her," which unfortunately included "obscene comic invention rushed to fill in the vacuum of accurate historical knowledge."[13] The earliest sources to describe her as homosexual were written during the Hellenistic period, with a recovered papyrus scroll from the late third or early second century

[11] Ibid

[12] Elizabeth Prettejohn, "Solomon, Swinburne, Sappho," Victorian Review 34, no. 2 (Fall 2008): 103.

[13] Glenn W. Most, "Reflecting Sappho," Bulletin of the Institute of Classical Studies, vol. 40, Oxford University Press (1995): 17.

B.C. accusing her of "being irregular in her ways and a woman-lover."[14] Even so, Early Modern translations heterosexualized her poems, or worked on the assumption that she had only male lovers. For instance, in his 1711 translation of the "Ode to Aphrodite", Ambrose Philips described Sappho's beloved as male, which set a basis for preceding translators until the 1900s.[15] Additionally, in 1781, Alessandro Verri interpreted fragment 31 as being about Sappho's love for a man, and Karl Otfried Müller wrote that the poem described "nothing but a friendly affection" between the women.[16]

Today, countless translators and scholars agree that many fragments of Sappho's surviving poetry "clearly" describe and celebrate "eros between women."[17] Mellissa Mueller even muses that the fragments can be read as queer even if the "question of her lesbianism is undecidable."[18]

Sappho by Solomon

Solomon created a graphite sketch of Sappho, dated April 14th, 1862 (**Figure 1**), which is one of his first classical depictions (if not his first). She is in profile, with a long, Roman-esque nose and a deep brow. She is rendered from the neck up, with a slight outline of her shoulder. Her hair is dark, bushy, and down. Here, Sappho is represented "unequivocally as a poet".[19]Her eyes are closed in contemplation, her lips slightly parted, and a laurel wreath, a symbol of the classical poet, sits atop her head. Colin Cruise, a scholar of Solomon, suggests two possibilities for the model of this sketch. Keomi Gray, a woman of Romany origin and frequent model for the PRB circle, is one possibility. Another is Jamaican model Fanny Eaton, also a commonly used muse of the Brotherhood. Both of these women had dark complexions, similarly to how Sappho's tradition characterizes her.[20]

---

[14] D.A. Campbell, "Greek Lyric 1: Sappho and Alcaeus," Loeb Classical Library no. 142, Harvard University Press (1982)

[15] Joan DeJean, Fictions of Sappho: 1546–1937," Chicago: University of Chicago Press (1989): 318.

[16] Most, "Reflecting Sappho," 26-28.

[17] Sandra Boehringer, "Female Homoeroticism," in A Companion to Greek and Roman Sexualities, Chichester: Wiley Blackwell (2014): 151.

[18] Melissa Mueller, "Sappho and Sexuality," In The Cambridge Companion to Sappho, Cambridge: Cambridge University Press (2021): 47.

[19] Elizabeth Prettejohn, "Solomon, Swinburne, Sappho," 107.

[20] Ibid, 108

Figure 1.
Solomon, Simeon, *Study of Sappho*, Pencil on Paper, 1862,
Tate, London, Simeon Solomon Research Archive

Two years later, Solomon created *Sappho and Erinna in a Garden at Mytilene* (**Figure 2**). Unlike the previously mentioned sketch, this piece was completed in beautifully bright watercolors. The figure of Sappho is centered, wearing a long, flowing pale mustard-yellow chiton, with bare feet. One knee is bent towards the other figure, the other leg stretches to the other side, as though mid-motion. Identical to the sketch, atop her raven head sits a green laurel wreath, a customary token of an ancient poet/poetess, or more broadly a master of the arts. Her face appears strikingly similar to the Sappho sketch: she is turned to her right, with an androgynous profile, long nose, deep brow, pouting lips, and round chin. There is no doubt that Solomon either used the same model or referenced his previous sketch. However, contrary to the sketch, the Sappho of the watercolor painting's emotions are clearer. While in pencil she appears thoughtfully soft, as if internally putting together her next poem, in watercolor she is anguished, with a deeply furrowed brow. Her face leans on the woman the title identifies as Erinna's cheek, her right arm wrapped around the other

Figure 2.
Solomon, Simeon, *Sappho and Erinna in a Garden at Mytilene,*
Watercolor, 1864, Tate, London, Simeon Solomon Research Archive

woman's back, long fingers resting on her waist, while her left crosses Erinna's front. Sappho clutches Erinna's shoulder, while the latter rests her hand on top of the former's. Erinna wears a chiton in a similar style to Sappho's, though in magenta. Her ginger hair contrasts Sappho's dark locks. Her eyes are large and looking directly at the viewer, inviting us in, and her features are softer and more feminine than the other woman's. Her breast is partially exposed, concealed by Sappho's upper arm. One knee rests against Sappho's in a sense of familiarity, and her other hand rests limply in her lap. Although her displeasure is not as evident as Sappho's, I argue that it is still clearly present. Her downturned doe eyes, her slightly pursed lips, her slumped shoulders, and the consoling hand atop Sappho's are all indications of a shared sense of a haunting grief.

Through the influence of the budding ideas in London, such as Neoclassicism, Swinburne became aware of the surviving fragments of the 6th Century BCE Greek poetess. In the 1860s, he wrote and released "Sapphics," which included the following:

While the tenth sang wonderful things they knew not.
Ah the tenth, the Lesbian! the nine were silent,
None endured the sound of her song for weeping;
Laurel by laurel,

Faded all their crowns; but about her forehead,
Round her woven tresses and ashen temples
White as dead snow, paler than grass in summer,
Ravaged with kisses,

Shone a light of fire as a crown for ever.

Yea, almost the implacable Aphrodite
Paused, and almost wept; such a song was that song.
Yea, by her name too

Called her, saying, “Turn to me, O my Sappho;”
Yet she turned her face from the Loves, she saw not
Tears for laughter darken immortal eyelids,
Heard not about her

Fearful fitful wings of the doves departing,
Saw not how the bosom of Aphrodite
Shook with weeping, saw not her shaken raiment,
Saw not her hands wrung;

Saw the Lesbians kissing across their smitten
Lutes with lips more sweet than the sound of lute-strings,
Mouth to mouth and hand upon hand, her chosen,
Fairer than all men;

Only saw the beautiful lips and fingers,
Full of songs and kisses and little whispers,
Full of music; only beheld among them
Soar, as a bird soars[21]

[21] Algernon Charles Swinburne, “Sapphics by Algernon Charles Swinburne,” Poetry Foundation, accessed May 3, 2024. https://www.poetryfoundation.org/poems/45302/sapphics56d224c13e1d5.

Swinburne reinterprets Sappho's "Ode to Aphrodite" in this poem, where he not only solidifies her homosexual relationships but also worships her as a poet. Solomon may have been inspired by this explicit worship, which may additionally explain his choice of another poetess as a subject to pair with Sappho. There is a clear reference to Sappho as the tenth muse, as Plato called her.[22] Laurel crowns, Aphrodite, doves, and instruments, which all make appearances in *Sappho and Erinna in a Garden at Mytilene*, are all mentioned in Swinburne's text. Additionally, Swinburne writes "mouth to mouth and hand upon hand" to describe the couple's actions, which correlates to how Sappho and Erinna are posed.

In Swinburne's poetry regarding Sappho, there is no mention of Erinna. Even in Sappho's own surviving poetry, the name Erinna does not come up. So why does Solomon pair them together? Erinna was an ancient poetess, perhaps the most well known after Sappho. A 10th century encyclopedia lists her as a companion of Sappho, but scholars now generally agree that she actually lived hundreds of years after her.[23] Perhaps Solomon operated on the assumption that they were companions, or perhaps by using the name Erinna, instead of a name of one of Sappho's many documented female lovers, he further makes this piece of art akin to a piece of poetry by pairing together two celebrated poetesses in such a loving way.

I propose that Solomon's inspiration came from not only Swinburne's poetry, but also directly from the poetess herself. Two fragments of Sappho's poetry are believed to be discussing her desire for another woman, Anactoria. In fragment 16, Sappho reminisces on her lover's beauty, and compares her departure to Helen of Troy's from Paris:

> Some say it's a force of cavalry, others of foot,
> others of ships, but I say that the most beautiful
> thing upon the black earth is whatever it is you
> desire.
>
> It's easy enough to make this plain to all:
> for she who was far more beautiful than any woman of mortal
> race,
> Helen, abandoned her husband - the best of men -

[22] Elizabeth Prettejohn, "Solomon, Swinburne, Sappho," 103.

[23] Donald Norman Levin, "Quaestiones Erinneanae," Harvard Studies in Classical Philology, Department of the Classics, Harvard University, 1962: 193.

and went sailing off to Troy; she remembered neither her
child nor her much-loved parents, but Aphrodite(?) led her
astray . . .

[This] has made me think of Anactoria, who isn't here.

Her step, which stirs desire, and the bright sparkle of her
face,
are dearer sights to me than the chariots of Lydia,
and armed men fighting on foot.
[24]

The grief that Sappho feels after her lover's departure can be read similarly to her dismay expressed in *Sappho and Erinna in a Garden at Mytilene*. Sappho clutches at Erinna's shoulder, wraps her arm around her waist, and is posed in a way that makes her legs appear as if they are either in motion or were just in motion. Contrarily, Erinna's limbs are idle. She looks at us, while Sappho looks at her. In my understanding, this piece could very well be directly inspired by fragment 16. The watercolor's title suggests nothing for the context of the aura of grief and desperation, but the fragment does. Sappho's lover is leaving her, and she is immensely disheartened. Sappho appears as the one being *left*, with her desperate clutch and sorrowful expression, while her lover is the one *leaving*, with her defeated pose and sad, perceptive eyes gazing at us. Although the name Erinna is in the title and not Anactoria, there seems to be direct influence on the piece from this fragment.

Sappho and Erinna in a Garden at Mytilene can also be read as referencing fragment 1, known as "Ode to Aphrodite", in which the poetess calls upon the goddess of love and beauty for romantic help.[25] Although the scenes do not directly match (neither does fragment 31, but the allusion is apparent), there are several references to Aphrodite in this piece. To the far mid-upper right there is a statuette, presumably of Aphrodite (due to her hair style, nude upper half, and hand covering her vaginal area resembling the classic pudica pose). The statuette reaches one hand out towards the couple. At her feet is a papyrus scroll, with illegible writing, and a myrtle branch laying in the roll of the paper. Myrtles, along with roses, are sacred to Aphrodite. Perhaps, then, the illegible writing depicts

[24] Sappho, "Fragment 16", trans. Gillian Spraggs, Love Shook My Senses. Lesbian Love Poems, The Women's Press, London (1998)

[25] Elizabeth Prettejohn, "Solomon, Swinburne, Sappho," 116.

fragment 1.[26] Roses appear throughout the painting, another direct attribute. The deer could be either a reference to Artemis (or Diana), whose sacred animal was a deer. Artemis led an all-female group, similarly to how Sappho led a group of poetesses.

The deer could also be in reference to Apollo, master of the Muses, as he shares the sacred animal with his twin sister. This would be alluding to Sappho's honorary title of the tenth muse. Two doves, the love goddess's sacred animal, sit directly above the two women, mirroring their pose. The left doves' body is partially turned towards the viewer, as is Erinna's, and the right dove is turned to the side, as is Sappho. The doves also sit directly above the women, clearly indicating a symbolic theme of love. A black crow sits beside the pair of doves, mouth open in a squawk. This could be referencing fragment 31, one of Sappho's most famous poems, where she expresses intense, physical jealousy upon seeing her lover Anactoria with her betrothed. The bothersome crow could be alluding to the unwanted presence of the male. The fragment goes as follows:

He seems to me the peer of gods, that
man who sits and faces you, close by you
hearing your sweet voice speaking,

and your sexy laugh, which just this moment
makes the heart quake in my breast: for every time
I briefly glance towards you, then I lose all power
of further speech.

My tongue is smashed; at once a film of fire
runs underneath my skin; no image shapes
before my eyes;
my ears are whining like a whirling top;

cold sweat pours down me, and in every part

shuddering grips me; I am paler than

[26] Ibid, 117-117

summer grass, and seem to myself to need

little to make me die.[27]

In 1865, Solomon produced *Erinna Taken from Sappho* (**Figure 3**), another sketch seemingly unfinished, which must be directly inspired by fragment 31. We can see references to Solomon's previous piece; a column with a statuette, a scroll, a myrtle branch, and a lyre – all symbols for both poetry and love. There is an obvious contrast between the figures of Erinna and her betrothed and Erinna and Sappho in the watercolor. While Sappho's Erinna is unexposed and seems to share her lover's grief, the Erinna of this sketch seems irritated by the man's grasp and looks down in sadness, her breast exposed. The earlier depiction of Erinna rests her hand atop Sappho's, while this Erinna appears as though she is in the act of pushing the man's hand away from her body. The two doves and blackbird as painted in *Sappho and Erinna in a Garden at Mytilene* seem to mirror

Figure 3.
Solomon, Simeon, *Erinna Taken from Sappho*, Pencil on Paper, 1865, Private Collection, Simeon Solomon Research Archive.

[27] Sappho, "Fragment 31", trans. Gillian Spraggs, Love Shook My Senses. Lesbian Love Poems, The Women's Press, London (1998)

the figures in the sketch. However, in the case of *Erinna Taken from Sappho*, Sappho becomes the squawking crow, perhaps symbolizing her agony over her lost lover.

Together, these three works of art are perhaps the most accurately rendered and breathtakingly alluring art pieces that show the ancient poetess how her direct word tells us she was. Solomon, being a homosexual, could create works of art with female subjects without the glaring presence of the male gaze, unlike many artists before him as well as his contemporaries. The feeling of queer longing and understanding is so apparent that it almost takes on physical presence within these works.

Classical Lesbians as Portrayed in Painting

There are two instances of what can be described as lesbian (appearing) love affairs in Ovid's *Metamorphoses*. Ovid's work has been an inspiration to artists for centuries, yet both of these stories have a small amount of artworks dedicated to them. Ovid expresses why this may be in book nine of *Metamorphoses*:

> Mares do not burn with love for mares or heifers for heifers: the ram inflames the ewe: its hind follows the stag. So, birds mate, and among all animals, not one female is attacked by lust for a female.[28]

As we can see, romantic relationships between two women were deemed impossible to men like Ovid. Sexual relationships between men, however, were normal and accepted – this fact alone shows us how ancient the patriarchy is.

This quote is derived from Ovid's story of Iphis and Ianthe, in which a man informs his pregnant wife that if she gives birth to a girl, they must kill the child. She does end up giving birth to a baby girl, Iphis, but raises the child as a boy in order to save her life. When Iphis is thirteen, her marriage is arranged to Ianthe, who is extremely happy about the arrangement. Iphis is also already in love with Ianthe, but is nervous about the marriage because she knows she can't perform her "husbandly" duty of mating with his wife. Iphis and her mother go to the temple of Isis, where they pray to the goddess for help. The aid comes in the form of a transformation—Iphis

[28] Jody Valentine, "Ovid: The Metamorphoses : Book IX: Lines 666 – 797," Gender and Sexuality in Ancient Rome, accessed April 2nd, 2024, https://pressbooks.claremont.edu/clas112pomonavalentine/chapter/ovid-themetamorphoses-book-ix-lines-666-797-the-english-translated-by-a-s-kline-

is transformed into a boy.[29] This Ovidian story is historically a less-popular source of inspiration for artists, but there are a handful of artworks done in mediums like etchings and prints. However, these artworks all depict the moment in which Iphis and her mother are begging for help, emphasizing the importance of the gender transformation.

Ovid's story regarding Diana and Callisto goes as follows: Jupiter (or Zeus, in Greek mythology), becomes infatuated with a young nymph, Callisto, who is a follower of Diana (or Artemis, for the Greeks). He wishes to have sex with her, and in order to do so while hiding from his famously jealous wife, he transforms into Diana. While in Diana's form, Jupiter seduces Callisto.[30] There are many renditions of this scene, as well as the events that follow in Ovid's story, depicted in Early Modern art. Specifically in the case of Jean Baptiste Marie Pierre's *Diana and Callisto*, François Boucher's *Jupiter, In the Guise of Diana*, and *Callisto*, and Nicolas-René Jollain's *Diana et Callisto*, there is an obvious sexualization of these females figures – because what's better than one overly-sexualized female body than two entangled together. But something else lies within. As the story goes, two *women* did not actually have sexual intimacy – instead yet another woman tricked into having sex with Jupiter.

Both these stories had one thing in common – intimacy between a man and a woman never actually took place. The relationships between women were never genuine. And although it is implied that Callisto was infatuated with Diana, and more than implied that Iphis, as a woman, was in love with Ianthe, these realities are dismissed. Instead, artists either emphasized that a man was in the relationship all along, or overly sexualized the two female bodies.

We can thank Solomon for doing neither; for not following in the footsteps of the only artwork depicting classical lesbian relationships he had to reference. Through his *Study of Sappho, Sappho and Erinna in a Garden at Mytilene*, and *Erinna Taken from Sappho*, Solomon showed the art world that genuine classical sapphic relationships deserved recognition.

Legacy

On February 11th, 1873, at thirty-two years old, Simeon Solomon was arrested after being found engaged in sexual acts with another man in a public urinal. After a six week imprisonment, he was released to a family

[29] Christine Downing, "Lesbian Mythology," Historical Reflections / Réflexions Historiques 20, no. 2, Special Issue: Lesbian Histories (Summer 1994): 170.

[30] Ovid, Metamorphoses, Book II (A. S. Kline's Version): lines 401-465.

member on a £100 bail. After his time in jail, he was admitted into two separate lunatic asylums, most likely by family members.[31] The following year in Paris, he was once again arrested for a similar act in a public urinal and served three months imprisoned. He was ostracized from the public art world after these events and unfortunately fell into alcoholism. Many of his colleagues left him behind, not wanting to be publicly associated with a now known "sodomite."[32]

It is devastating that he passed away ostracized by the art world in which he held so dear. Fortunately, through the increasing popularity of queer studies, Solomon is now regarded as a queer icon, and rightfully so. He made choices in his career that some are still afraid to make today and never backed away from his queer identity. We can thank him for creating these pieces of art that are so different from previous artworks that attempt to show instances of classical lesbian love. His understanding of queer identity is clear in these artworks, and with the absence of the male gaze, he gave us beautiful depictions of the oldest documented lesbian woman and helped to give her the legacy she deserves.

[31] Carolyn Conroy, "Solomon's Life After 1873," Simeon Solomon Research Archive, accessed April 19, 2024, https://www.simeonsolomon.com/.

[32] Ibid.

Bibliography

Boehringer, Sandra. "Female Homoeroticism." A Companion to Greek and Roman Sexualities, Chichester: Wiley Blackwell (2014).

Conroy, Carolyn. "Solomon's Life After 1873." Simeon Solomon Research Archive, accessed April 19, 2024. https://www.simeonsolomon.com/.

DeJean, Joan. "Fictions of Sappho: 1546–1937." Chicago: University of Chicago Press (1989). Downing, Christine. "Lesbian Mythology." Historical Reflections / Réflexions Historiques 20, no. 2, Special Issue: Lesbian Histories (Summer 1994).

Easby, Rebecca Jeffrey. "A beginner's guide to the Pre-Raphaelites." *Smarthistory*, last modified August 9, 2015. https://smarthistory.org/a-beginners-guide-to-the-pre-raphaelites/.

Easby, Rebecca Jeffrey. "The Aesthetic Movement." *Smarthistory*, last modified June 3, 2016. https://smarthistory.org/the-aesthetic-movement/

Eaton, A.W. "Feminist Philosophy of Art." Philosophy Compass 3, no. 5 (September 2008). Elizabeth Prettejohn, "Solomon, Swinburne, Sappho." Victorian Review 34, no. 2 (Fall 2008).

Ferrari, Roberto. "Solomon's Life Before 1873." *Simeon Solomon Research Archive*, accessed April 19, 2024. https://www.simeonsolomon.com/.

Glenn W. Most. "Reflecting Sappho," Bulletin of the Institute of Classical Studies, vol. 40, Oxford University Press (1995).

Hamilton, Richard, and D. A. Campbell. "Greek Lyric i: Sappho and Alcaeus." *The Classical World* 77, no. 1 (1983). https://doi.org/10.2307/4349505.

Levin, Donald Norman. "Quaestiones Erinneanae." Harvard Studies in Classical Philology, Department of the Classics, Harvard University, 1962.

Mueller, Melissa. "Sappho and Sexuality." The Cambridge Companion to Sappho, Cambridge: Cambridge University Press (2021).

Ovid. Metamorphoses, Book II (A.S. Kline's Version).

Sappho. “Fragment 16”, trans. Gillian Spraggs. Love Shook My Senses. Lesbian Love Poems, The Women’s Press, London (1998).

Sappho. “Fragment 31”, trans. Gillian Spraggs. Love Shook My Senses. Lesbian Love Poems, The Women’s Press, London (1998).

Swinburne, Algernon Charles. “Sapphics by Algernon Charles Swinburne,” Poetry Foundation, accessed May 3, 2024. https://www.poetryfoundation.org/poems/45302/sapphics-56d224c13e1d5.

Valentine, Jody. “Ovid: The Metamorphoses: Book IX: Lines 666 – 797.” Gender and Sexuality in Ancient Rome, accessed April 2nd, 2024. https://pressbooks.claremont.edu/clas112pomonavalentine/chapter/ovid-the-metamorphoses-book-ix-lines-666-797-theengl ish-translated-by-a-s-kline-copyright-2000-all-rights-reserved/.

Ward, A.W., G.W. Prothero, and Stanley Leathes. “The Classical Renaissance.” The Cambridge Modern History 1, Cambridge University Press (1907).

Addison Cucchiaro is currently pursuing a bachelor's degree in Art History with a minor in Italian Studies at the University of Memphis. She is on track to graduate in May 2025 and will be attending the graduate program in Art History at American University in Washington, D.C. this fall. Addison's long-term goal is to earn a PhD in Art History, with a focus on the Italian Renaissance and Baroque periods. Addison's decision to submit her work to *QuaesitUM* was motivated by a desire to share her academic findings with a wider audience. Her research, particularly in the fields of gender and queer studies in art history, highlights the intersections of art and identity, exploring the ways in which gender and sexuality have been represented in art history. Her research aims to bring forward these underrepresented perspectives and challenge traditional narratives.

Addison Cucchiaro

The Sleeping Beauty: *Venus and Cupid* by Artemisia Gentileschi and the Spectacle of a Goddes at Rest

Faculty Sponsor
Dr. Rebecca Howard

Abstract

This research explores the transformation of the depiction of feminine divinity in Western art, focusing on the shift from powerful mythological goddesses to passive, objectified female figures under the male gaze during the Renaissance and Baroque periods. Through the analysis of works by male artists, the research examines how the motif of the sleeping goddess reflects the growing influence of patriarchy in early modern Europe. The analysis then shifts to Artemisia Gentileschi's *Venus and Cupid*, which challenges these conventional representations by presenting Venus as a serene, empowered figure, free from the constraints of male spectatorship. Comparing Gentileschi's work to that of her male counterparts highlights the differences in how male and female artists presented their subjects, down to the smallest details. Gentileschi's interpretation of a sleeping Venus reclaims the image of the divine feminine, providing a more empowering understanding of women and their roles in both art and spirituality. Through a feminist lens, this research argues that Gentileschi's work offers a critical reimagining of feminine power, presenting a divine femininity that is not bound by objectification or the male gaze.

Introduction

The depiction of feminine divinity has undergone profound transformations throughout history, reflecting broader cultural and societal shifts. From ancient statues of pagan goddesses to the idealized nudes of early modern European art, the portrayal of women in divine or mythological contexts reveals the growing influence of the patriarchy in art. While figures like Venus, the Roman goddess of love and beauty, have embodied fluctuating symbolism across time and space—sometimes representing feminine power and fertility, at other times serving as an object for male pleasure—male artists of the Renaissance and Baroque periods often emphasized the latter. Their depictions reinforced Venus's perceived passivity, stripping her of agency and rendering her a vulnerable, voyeuristic spectacle, governed by male-dominated artistic ideals. However, Artemisia Gentileschi's *Venus and Cupid* (**Figure 1**) challenges this narrative, reclaiming the image of Venus as a figure of serenity, empowerment, and maternal connection, free from the constraints of the male gaze. Through an analysis of historical contexts, depictions of the sleeping goddess by male artists, and Artemisia Gentileschi's work, we can explore early modern perceptions of gender, highlighting how artistic choices and feminist perspectives differ between male and female artists.

Figure 1.
Artemisia Gentileschi, *Venus and Cupid*, c. 1625-26, oil on canvas,
Virginia Museum of Fine Arts

Analyzing the Sleeping Goddess

Scholars suggest that the suppression of goddess worship coincided with the rise of Christianity. However, the fall of paganism could not "subdue the impulse to worship female incarnations of the divine."[1] Many local goddesses were transformed into saints, and Christian churches often replaced pagan shrines. A notable example is the Church of Santa Maria sopra Minerva in Rome, built over a temple dedicated to the Greek goddess Athena.[2] This continuity underscores the enduring need for feminine representations of divinity. This instinct is additionally seen in the popularity of the worship of Mary as the Mother of God. For many worshipers, Mary's status became equal to that of Christ himself, with an emphasis on her role as the Mother. In a world increasingly shaped by patriarchy, childbirth was one of the few powers exclusively held by women. Men relied on women to perpetuate life, granting them an indispensable role in the continuation of society. In this way, all women, through their ability to bear children, wielded a unique and undeniable form of power.[3]

The image of a divine feminine figure plays a crucial role in shaping-self-perception. Without a feminine likeness of God to relate to, women risk being relegated to a lesser status in both spiritual and societal hierarchies. Divine representations are therefore crucial—not only for spiritual equality but also for affirming women's inherent worth and power.[4]

The Roman goddess Venus has been a central figure in art for centuries, serving as one of the most enduring and influential muses in early modern art history. From the decline of pagan worship to the creation of *Sleeping Venus* (**Figure 2**) by Italian Renaissance artist Giorgione (born Giorgio Barbarelli da Castelfranco), the passage of time and the rise of patriarchy drastically transformed how this goddess, long associated with female sexuality and womanhood, was depicted.[5]

---

[1] Judy Chicago and Edward Lucie-Smith, *Women and art: Contested territory*, (Hertfordshire: Eagle Editions, 2004), 24.

[2] Ibid

[3] I note that not all women have a uterus, that not all people with a uterus identity as a woman, and that not all women can/want to bear children. I use the term women to describe those who were assigned female at birth and maintain that gender identity, keeping in mind the typical standards of gender within the periods of time described.

[4] Chicago and Lucie-Smith, *Women and art: Contested territory,* 20.

[5] *Venus Sleeping* is traditionally attributed to Giorgione, however it has been thought that Titian completed the painting following Giorgione's death. Both the landscape and sky are overall accepted to be primarily done by Titian's hand.

Figure 2.
Giorgione, Titian, *Sleeping Venus*, c. 1510, oil on canvas,
Gemäldegalerie Alte Meister, Dresden

Sleeping Venus is recognized as the first known reclining female nude in modern Western art history, and established a genre of similar subjects paired with landscapes. When the female nude was initially popularized in early modern art, mythological women were mainly used as muses. This allowed the artist to "distance the image from the spectator so that, while the nude was apparently sexually available, it did not become sexually threatening."[6] These women were otherworldly and therefore available to be rendered in ways that real, living women were not socially and morally allowed to be.

Giorgione's painting is a clear example of the male gaze as manifested in the art world. "The male gaze," Mary Devereaux explains, "involves more than simply looking; it carries with it the threat of action and possession." The power to possess is not reciprocal.[7]

---

[6] Chicago and Lucie-Smith, *Women and art: Contested territory*, 102.

[7] Mary Devereaux, "Oppressive Texts, Resisting Readers, and the Gendered Spectator," (*Feminism and Tradition in Aesthetics*, 1995)

Giorgione's Venus lays on the ground, reclining atop a silvery-white textile that contrasts against a deep crimson pillow supporting her head, creating a striking balance of light and shadow. Her body is positioned diagonally across the canvas, drawing the viewer's eye along the curve of her figure. One arm is raised and bent behind her head, granting access to a view of her breasts. The other hand rests suggestively at her groin, with her fingers curled inwards. Venus's body is depicted as unblemished and smooth, with a flawless, almost ethereal quality. She obviously sleeps, as the title suggests, as her eyes are shut and her features are soft. The background of the painting features a quintessential Venetian-inspired landscape, with rolling hills, distant mountains, lush trees, calm waters, and quaint buildings scattered throughout. The landscape's fertile, abundant qualities could symbolize Venus's role as a goddess of fertility. Little is known about the meaning behind the painting or its relationship to other representations of sleeping women. Scholars have discovered that the work had been commissioned by Girolamo Marcello as a wedding present.[8] Images of Venus, often nude, were often commissioned for weddings. They were meant to both visually please the groom and visually inspire the bride.

In 1603, Italian artist Annibale Carracci produced *Sleeping Venus* (or *Sleeping Venus with Putti*) (**Figure 3**). Here, Venus lays in a lavish bed placed outdoors, surrounded by tens of putti. Similarly to Giorgione's *Venus,* her arm is raised and bent behind her head. Her other hand is not directly on top of her groin, but it does rest between her legs. The title may suggest that the goddess is sleeping, but a quick visual analysis shows that she could be described more fittingly as *posing*. The fingers of her right hand are flexed straight, as are her toes. Everything about her pose is both uncomfortable and unnatural. Around her, putti talk to one another, climb trees, play instruments, and overall act like rowdy babies. Yet, as the title insists, she sleeps.

[8] Udo Kultermann, "Woman Asleep and the Artist," (*Artibus et Historiae*, Vol. 11, No. 22 1990), 136.

Figure 3.
Annibale Carracci, *Sleeping Venus*, c. 1603, oil on canvas,
Musée Condé, Chantilly, Oise, France

In the 1528 painting *Venus and Cupid with a Satyr* (**Figure 4**) by Antonio da Correggio, the sleeping Venus is depicted as a vulnerable and passive figure, her body is exposed to both the viewer's gaze and the male figure within the composition. The adult male figure, a satyr, observes her with desire. Her slumber enhances her vulnerability, as she remains unaware of their watchful eyes. Similarly, Nicolas Poussin's *Sleeping Venus and Cupid* (1630) (**Figure 5**) features the passive figure of Venus lying in slumber, while two men in the background gaze at her from behind a tree, adding an element of voyeurism.

The passive, dormant Venus is positioned as an object of desire, subject to the lustful gaze of the male figures. One of her arms is raised, and her legs are positioned apart. This depiction underscores the theme of the male gaze, wherein Venus is reduced to a mere object for male pleasure.

Figure 4.
Antonio da Correggio, *Venus and Cupid with a Satyr*, c. 1528,
oil on canvas, Musée du Louvre

Figure 5.
Nicolas Poussin, *Sleeping Venus with Cupid*, c. 1630,
oil on canvas, Gemäldegalerie Alte Meister, Dresden

Discussion

Why did these artists choose to depict Venus in a state of sleep? In their paintings, Venus's sleep, coupled with her nudity, renders her completely vulnerable—unable to resist or even be aware of the gaze of the spectator. She exists to be a "passive object of the male libido."[9] She becomes entirely passive, a figure who exists solely for the male viewer. This portrayal transforms Venus, once a powerful symbol of feminine divinity and sovereignty, into an object for male desire. By depicting her as asleep, the male artists turn Venus, traditionally seen as the embodiment of the divine feminine power, into a passive sex symbol. The innate power of the female body was overpowered by the priority of a man's pleasure. Venus – now the perceived epitome of sex and the idealized beauty – is put to sleep. Why is such a historically divine figure like Venus, one of the most respected and worshipped goddesses of antiquity, depicted in such a vulnerable way? Why is she asleep, and why is she watched?

The act of putting Venus to sleep serves as a powerful symbolic gesture. Sleep is often associated with passivity, submission, and an unconscious state, making it a fitting metaphor for the way in which these artists depict Venus as no longer an active, powerful deity but rather an object. Her nudity, a symbol of beauty and fertility in classical mythology, is now stripped of its divine power and reduced to a visual delight for the male spectator.

By painting the classical personification of love, beauty, and sex in a vulnerable state, the artist is telling his audience that they, as men, have control over such things. If they can possess a goddess in such a way, they can possess real women in such a way. It is essential to challenge the traditional, male-centered interpretations of how sleeping women have been portrayed in art and to reevaluate how these depictions reflect broader-gender dynamics in society. The fetishization of the nude female sleeping without religious or ritualistic function could have originated as far back as Roman times. In the first century B.C., Roman poet Propertius, in his third Elegy, described a slumbering woman as viewed by her lover. In this poem, Propertius "demonstrates a male perspective towards a woman reduced to an object." The nature of their relationship is unknown, and she is

[9] Chicago and Lucie-Smith, Women and art: Contested territory, 101.

unable to respond to his gaze. An image like the one described by Propertius paralleled the role of women in Roman society, which was "determined by the emotions, reactions, and exploitation of women."[10]

Works by artists such as Giorgione, Carracci, Correggio, and Poussin demonstrate how the motif of the sleeping mythological woman became a favored theme during the Renaissance and Baroque periods. Paintings of this nature can undoubtedly be described as the epitome of the male gaze as seen in early modern art, but what happens when the gaze shifts? What happens when the person holding the paintbrush is a woman?

Venus and Cupid (also known as *Sleeping Venus*) (**Figure 1**) was made around 1625 - 1626 by Italian Baroque artist Artemisia Gentileschi. The goddess lays on a plush bed of a rich blue color, which was painted with two layers of lapis lazuli.[11] Her head rests on a red pillow with gold tassels. The expertly rendered fabrics undoubtedly were inspired by Orazio Gentileschi, her father and fellow Baroque artist who famously excelled in fabrics.

The deep red color of both the pillow and curtains behind her could be influenced by Caravaggio, who worked with Orazio. She wears nothing but a thin strip of sheer fabric that encircles her right thigh and weaves behind her back and around her left upper arm, ending at her hand. Her fingers loosely clasp the end of the wispy scarf. Her other hands rest gently on her waist. Her milky-white skin almost appears as if it is glowing. Venus' body is extremely naturalistic, just as is every other nude female body painted by Artemisia. Her curves, her stomach, and her breasts appear almost indistinguishable from a true bare female form. Her round cheeks, long nose, and small chin are all facial features of Artemisia herself. It was common for the artist to use herself as a muse. The face is in profile, facing up towards the heavens, not towards the presumed viewer (like Giorgione paints). There is a small, soft smile of ease and relaxation on her lips. On her ear we can see a large pearl earring, an attribute to Venus in reference to her birth from the sea.

Cupid, son of Venus, stands by his mother's head and peers down at her with a sense of familiarity. He appears a bit older than his usual depic-

---

[10] Udo Kultermann, "Woman Asleep and the Artist," (*Artibus et Historiae*, Vol. 11, No. 22 1990), 130-134

[11] Keith Christiansen and Judith W. Mann, *Orazio and Artemisia Gentileschi* (New York: The Metropolitan Museum of Art, 2002), 372.

tion as an infant, perhaps being an older toddler. Like his mother, he wears only a thin, wispy strip of fabric. His typical wings are alert and perked. In one hand, he holds a bundle of peacock feathers. In antiquity, peacock feathers were usually attributed to the queen of the gods and patroness and women and children, Juno. Although not an attribute to Venus or Cupid, it does make sense to place the feathers in such a setting; not only did they represent a goddess' royalty (and therefore her divinity), but also womanhood and motherhood. It is unclear exactly what Cupid is doing with the fan of feathers, but he could be in the act of fanning air on his mother or swatting away bothersome pests.

Unlike paintings made by her male counterparts, Artemisia's Venus is indoors and enclosed in her own space. A glimpse to the outside world can be seen through a window that takes up the upper left corner of the artwork. I propose that this can be read as Artemisia gifting Venus a privacy that her male counterparts do not. The lone temple among the flourishing landscape recalls ancient buildings dedicated to Vesta, Roman goddess of the home and family. Artemisia may have known this and may have seen such structures while working in Rome. Perhaps, she intended to set a pagan mindset and suggest "that the viewer participates in surreptitious worship of the goddess of beauty."[12] This would be highlighting the act of worshiping a female deity, and therefore retaining the idea of Venus as a divine goddess. While Artemisia utilizes common motifs found in paintings of the same subjects, such as the presence of Cupid, a visible landscape, and the act of sleep, the "emphasis of the painting is directed in new ways towards the viewer."[13]

Some scholars suggest that *Venus and Cupid* was commissioned, primarily due to the use of the expensive lapis lazuli and the unusual subject. However, no evidence of a patron survives, and at the time of this work's creation Artemisia was already very successful and could have had access to expensive supplies. It is true that the subject matter of *Venus and Cupid* may seem odd among Artemisia's portfolio. For instance, she is known for her paintings of heroines (such as *Judith in Judith Slaying Holofernes* and *Jael in Jael and Sisera*), of victims (such as Lucretia in *Lucretia* and Danaë in *Danaë*), or figures that overlap as both heroine and victim (such as Susanna in *Susanna and the Elders and Cleopatra in Cleopatra*). As a teenager, Artemisia was raped by her tutor, who was also a friend of her

[12] Christiansen and Mann, *Orazio and Artemisia Gentileschi*, 373.

[13] Kultermann, "Woman Asleep and the Artist," 140.

father. Around a year later, Orazio began the prosecution against his daughter's rapist. During trial, Artemisia had to endure not only reliving her trauma, but also torture by means of thumbscrews.[14] Her internal turmoil over these events is evident in her portfolio. Elizabeth S. Cohen proposes that Artemisia's history as a victim of rape "overshadows much discussion of the painter and has come to distort our vision," and that "biographers and critics have had trouble seeing beyond the rape."[15] Cohen implores that Artemisia's "reputation continues to be violated in the present by an overly sexualized interpretation."[16] This tells us why some scholars insist on a patron for the piece, and say that *Venus and Cupid* serves as evidence of Artemisia's "willingness to respond to the requests of male patrons."[17] I must disagree.

Even if this piece was commissioned, which there is no evidence of, I propose that Artemisia would not bend to the will of a male patron seeking a voyeuristic image of the Roman goddess. Several of her pieces of heroines and victims had patrons, and this did not yield her from using the artworks as outlets for her personal sentiments. At this point in her life, Artemisia was living as a single woman, after her husband had disappeared from her life in 1622, leaving her with her only surviving daughter (most of her children had died young).[18] She was a young, single woman, as well as a single mother. Her late adolescence had been disrupted by the sexual assault, but in the mid-1620s while she was in her early 30's, living without a husband and with a doting child, *Venus and Cupid* does not seem so out of place. Although Venus is no tyrant-slaying Judith or grief-stricken Lucretia, she is both a victor and victim in her own right. And, just as Artemisia paints Judith's features after her own, she paints her face on Venus. Unlike Venus as painted by Giorgione or Carracci, Artemisia offers a view of the sleeping goddess in an exclusively feminine, understanding way. *Venus and Cupid* shows the goddess in a genuinely restful state, not appearing as if posing for a voyeur. She is comfortable in her nudity and knowledgeable of her divine beauty. Her feminine body is celebrated, not fetishized. Her child brings a tender presence, as he dotes on his mother and reassures her comfort.

---

[14] Elizabeth S. Cohen, "The Trials of Artemisia Gentileschi: A Rape as History," (*The Sixteenth Century Journal, Special Edition: Gender in Early Modern Europe*, Vol. 31, No. 1, 2000), 47

[15] Ibid

[16] Ibid

[17] Christiansen and Mann, *Orazio and Artemisia Gentileschi*, 373.

[18] Cohen, "The Trials of Artemisia Gentileschi: A Rape as History," 50.

Conclusion

Artemisia Gentileschi offers us a rare and empowering vision of the divine feminine, presenting a figure of peace and serenity that is free from the oppressive gaze of male spectators. In her works, the absence of the male gaze allows women to encounter a reflection of the divine in themselves, portraying female figures not as passive objects but as subjects with inherent power, grace, and autonomy. Artemisia's art challenges traditional representations and enables women to see themselves as divine beings, capable of embodying strength and beauty without the distortion of male desire or objectification. This shift in perspective fosters a sense of spiritual equality, where the divine is not only reserved for men but also accessible to women.

Bibliography

Chicago, Judy, and Edward Lucie-Smith. *Women and Art: Contested Territory*. Hertfordshire: Eagle, 2004.

Christiansen, Keith, and Judith W. Mann. *Orazio and Artemisia Gentileschi*. New York: The Metropolitan Museum of Art, 2002.

Cohen, Elizabeth S. "The Trials of Artemisia Gentileschi: A Rape as History." *The Sixteenth Century Journal*, Special Edition: Gender in Early Modern Europe 31, no. 1 (2000).

Devereaux, Mary. "Oppressive Texts, Resisting Readers, and the Gendered Spectator." In *Feminism and Tradition in Aesthetics*, 1995

Kultermann, Udo. "Woman Asleep and the Artist." *Artibus et Historiae* 11, no. 22 (1990).

Jackson Lyles graduated *summa cum laude* and Honors with Thesis designation from the University of Memphis in May 2024. He earned his undergraduate degrees in international studies and nonprofit development and administration. During his undergraduate studies, he worked extensively on his honors thesis with Dr. Nicole Detraz and presented his preliminary research at the University Libraries' "NedTalks" event. Jackson chose his research topic of international nongovernmental organizations and social sustainability out of a deep interest in international studies and intercultural relations. Jackson has a strong desire to pursue a full-time career in nonprofit work focused on empowering, bringing justice to, and supporting the marginalized. He hopes to continue pursuing these interests through a graduate degree in City and Regional Planning at the University of Memphis.

Jackson Lyles

“Going Upstream”: Understanding Social Sustainability in an NGO Context

Faculty Sponsor
Dr. Nicole Detraz

Abstract

When it comes to humanitarian aid and development, International non-governmental organizations (INGOs) are significant actors on the international stage. Many of these organizations hold significant power as many of them have budgets larger than the countries they work in. This study investigates the challenges of social sustainability (SS) of various Northern INGOs. INGOs are forced to deal with relational power dynamics between their donors and beneficiaries. These dynamics have a profound impact on the SS of INGOs. This study is a qualitative content analysis of the websites of eight of the largest humanitarian and development INGOs. By analyzing text on these sites, this study is aimed at understanding how INGOs themselves conceptualize SS and their relational power dynamics with beneficiaries and donors. The study finds that SS is primarily conceptualized through the term "accountability" and that the INGOs are actively discussing SS. The paper finds that empirical data supporting INGO SS is sparse and concludes by suggesting that INGOs emphasize the publication of inward-looking documents, with supporting empirical data, in order to establish stronger SS practices.

Introduction

In November of 2021 the government of Mali banned France-funded aid nongovernmental organizations (NGOs) from operating in Mali (Ahmed, 2022). Why would Mali's government do this? Wouldn't Mali welcome such NGOs given the harsh realities many of its citizens face? After all Mali is widely considered as one of the poorest countries in the world—hundreds of thousands of Malians were receiving aid (Ahmed, 2022). France-funded NGOs in Mali were receiving 100 million euros from the French government from 2013 up until the ban (Ahmed, 2022). Mali's decision can be explained best by the long history of bitter French colonialism in Mali and throughout the region of West Africa coupled with the power dynamics between the two states (France considered as belonging to the global North while Mali belonging to the global South).

From 2009-2014 twelve major humanitarian international nongovernmental organizations (INGOs) lost over $2 million annually (Esslemont, 2015). One of these INGOs, World Vision International, reported that $1 million went missing from 2009-2013 (Esslemont, 2015). World Vision International discovered that this amount was lost due to corruption between staff and "outside vendors and bankers" as well as corruption among "internal staff" (Esslemont, 2015).

What do these two examples have in common? First, these examples highlight INGOs and how crucial they are in the humanitarian aid and development landscape. For decades, INGOs have been at the forefront of delivering humanitarian aid and development to communities in need (Gibson, 2019; Lindenberg & Doble, 1999; Unerman & O'Dwyer, 2010). After several years of declining world poverty rates, there have been increases in poverty caused by the many effects of the Covid-19 pandemic (Sánchez-Páramo et al., 2021). Even in the example of Mali, the very fact that French NGOs were banned shows that governments regard NGOs as having a significant role in the humanitarian aid landscape.

The need for effective INGOs in today's world is ever important. When it comes to poverty and vulnerable communities, INGOs are often the most equipped to combat poverty and to aid in community development efforts. Some member countries of the Organization for Economic Co-operation and Development (OECD) reported in 2008 that they provided $121.5 billion in "development aid" to developing countries (Unerman & O'Dwyer, 2010, p.477). Of this $121.5 billion sum, OECD countries gave $2.5 billion directly to nongovernmental NGOs (Unerman & O'Dwyer, 2010). This amount excludes that of private donations. In

2009, OECD countries reported that private donations from their citizens to NGOs totaled $23.8 billion. In these statistics, NGOs are defined as an organization that is "neither a commercial organization nor a public sector body" (Unerman & O'Dwyer, 2010 p.476). NGOs in these findings were also defined as those specifically working in "welfare and development" (Unerman & O'Dwyer, 2010 p.476). Given the large funds INGOs deal with, there is no doubt that INGOs are of vital importance when it comes to development. These statistics reveal what societies, at least those in the global North, view as a main solution to eliminating poverty—the INGO.

Second, these examples and statistics also highlight the extreme complexity of INGO relationships with their donors. INGO donors range from governments (like the previous example of the Mali government) to intergovernmental organizations such as the United Nations, to corporations, to private donors. Governments in particular have unique and complex relationships with INGOs that will in part be explained later in this paper. Governments often look to INGOs to complement and aid in social services. Consequently, many governments are donors to INGOs. INGOs regularly face pressure from such donors to offer effective services that satisfy their desires.

At the same time, INGOs must regularly respond to and serve their beneficiaries in the best way possible. Ideally, this should be an easy task for humanitarian INGOs since after all, the sole reason they exist is because poverty and numerous human injustices persist in the world today. However, in reality, this is nothing near an easy task. A group of researchers focusing on certain NGOs in Ghana and Uganda found that when these NGOs arrived in their respective communities, community members were left worse off than before the NGOs arrived (Gilbert, 2020). This was due primarily to the phenomenon of "crowding out" which is a process whereby preexisting government programs lose staffing and resources because incoming NGOs offer more attractive employee benefits. This results in an exodus of government employees to NGOs (Gilbert, 2020). The net effect of this process is harmful to the communities being served. Instead of complementing the existing government programs, the NGOs in the study reduced overall services to the populations in need by "out-competing" them (Gilbert, 2020).

A core part of this paper is the donor-INGO-beneficiary relationships like the ones highlighted in the examples above. These relationships evidently have a significant influence on the ability of INGOs to have a positive impact on the communities that they work with. The donor-INGO-beneficiary relationships are characterized by complex power dynam-

ics which directly affect the social sustainability of INGOs. This idea is reflected in the study, explained above, of NGOs in Ghana and Uganda.

With the donor-INGO-beneficiary relational power dynamics in mind, the central concept this paper seeks to examine is the topic of social sustainability. The United Nations was one of the first to coin the term "social sustainability". The UN states that social sustainability is "about identifying and managing business impacts, both positive and negative, on people" (United Nations Global Compact, n.d.). While the UN focuses mostly on social sustainability in the context of the for-profit sector, INGOs must navigate this concept arguably even more than for-profits do. This is because of the serious social aspects of the work that virtually every humanitarian INGO does. Just like for-profit companies, INGOs must evaluate their organizational structure, services, and even method of delivering services when thinking about effective social sustainability because of the unintended harm these things can cause to their beneficiaries.

Late human rights activist and Nobel Peace Prize winner, Desmond Tutu, once said, "There comes a point where we need to stop just pulling people out of the river. We need to go upstream and find out why they are falling in." This quote encompasses the entire reason for the writing of this paper. The real-world examples provided in this introduction convey both the importance of effective INGO work in vulnerable communities as well as the harmful effects of poor social sustainability. It is a simple fact that when INGOs enact poor social sustainability practices, vulnerable communities are harmed. There is a fine line INGOs must face between the pressures of their donors and the needs of the communities they serve. This tension affects the social sustainability of INGOs. Are INGOs discussing social sustainability? How do INGOs understand, and communicate concepts of social sustainability to the public? These are just some of the questions that form the basis of my research question: How do INGOs conceptualize social sustainability and the relational power dynamics with donors and beneficiaries?

For the purpose of this paper, humanitarian INGOs are defined as NGOs whose purpose is to provide humanitarian aid, relief, and development services to vulnerable communities globally. This definition of INGOs is similar to that of Unerman & O'Dwyer (2010). A difference between NGOs and INGOs is solely geographical. INGOs operate across nations and often have offices across the globe. This paper is focused solely on INGOs in the context of development and humanitarian aid though INGOs can have a wide range of functions. Generally, INGOs and NGOs, can be pursuing the same types of work. The only difference is that one

works internationally while the other does not. Further explanation of what INGOs are, how they are defined, and how the INGOs of study are selected in this paper are explained in the "methodology section".

This paper uses the UN definition of social sustainability although the definition of this term will be explained further in the next section. The paper will answer the research question using a systems thinking approach. A main idea behind this paper is that by defining the donor-INGO-beneficiary relationship as a system according to a systems thinking approach, the effect of this relationship on INGO social sustainability can be understood in a more clear and understandable way. Using this approach will help conceptualize the challenges INGOs face when it comes to social sustainability in ways that past research on this subject has not. There is no shortage of research regarding social sustainability and the donor-INGO-beneficiary relationship. However, there is not much research that examines how INGOs themselves are discussing and presenting these concepts. This paper seeks to fill in this gap.

What do I Mean by Sustainability?

In current scholarship, there appears to be a lack of an overarching term that encapsulates many of the terms I explained in the literature review. Accountability and its various types (O'Leary, 2017; Unerman & O'Dwyer, 2010), and human rights-based approaches (D'Hollander et al., 2013; Unerman & O'Dwyer, 2010) for example, have garnered a lot of attention in existing literature. It seems like terms like these are, in a way, getting at the same core challenge of how INGOs manage their impacts in the communities they serve in ethical ways. Yet, there is no overarching term to simplify the work in existing literature.

Therefore, I argue that the term social sustainability is an "umbrella" term, and the concept of accountability for example, can be situated under the umbrella of social sustainability. I will argue that accountability is the most commonly used term to refer to social sustainability among scholars and INGOs alike. One possible reason why social sustainability does not appear much in existing literature is because of its very broad connotations. The term can be interpreted and viewed in numerous ways and in even more numerous contexts. With this in mind, many of the sources I have cited do not specifically mention the term social sustainability but instead refer to various terms under, what I would argue, the "umbrella" of social sustainability. When social sustainability is mentioned in this paper, I am mainly referring to the terms under the umbrella concept. By

conceptualizing the term "social sustainability" as an umbrella term, I aim to simplify existing literature on the topic.

Who are Beneficiaries?

The term "beneficiary" is also vulnerable to wide ambiguity. It is clear that in existing literature on humanitarian and development INGOs that "beneficiary" is a very common term used to refer to any individual who is directly served and targeted by an INGO and is one who receives or takes part in an INGO program, product, or service (Bennett, 2010; Hughes, 2004; O'Leary, 2017; Unerman & O'Dwyer, 2010).

It is important to note that the term "beneficiary" implies that there is no action taken on the part of the receiver to receive the good or service being offered. This issue with the term beneficiary is highlighted in various scholarly works (Bhati, 2021; Gibson, 2019). Some authors go so far as to abandon the use of the term altogether (Gibson, 2019). The connotation of "beneficiary" goes in contradiction to the actual state of beneficiaries. In reality, and as seen in various pieces of literature, beneficiaries often work and labor for various goods or services they receive from INGOs (Krause, 2014). Despite the vulnerabilities to the term "beneficiary" I will use this term in my paper because it is a common term used in literature on this topic. Using this term will help others comprehend my research within the context of existing literature. It is not the scope of this paper to fully dismantle the term "beneficiary". A whole new research paper will likely be required to explore the use of this term.

Who are Donors?

This paper recognizes five types of donors within an INGO context. These five types are my own classifications, but they are primarily adopted from Bennett (2010), Brinkerhoff & Brinkerhoff (2004), and Jordan & van Tuijl (2006). For example, Bennett (2010) discusses government donors, as well as private donors including individuals and foundations, all of which are included in my definitions of donors. Brinkerhoff and Brinkerhoff, (2004), discuss intergovernmental donors such as the UN, World Bank, and USAID.

The first type of donor is private donors. These donors are individuals who donate their own personal money to the cause of INGOs. The second type of donor is a corporate donor. Corporate donors are businesses who donate a portion of their profits to an INGO. The third type of donor this paper recognizes is private and public charities and foundations. These types of donors are usually NGOs themselves, but mainly focus on pro-

viding grants and funds for NGOs they support. Intergovernmental organizations are the fourth type of donor. Intergovernmental organizations are global organizations that involve a coalition of governments. The United Nations and the World Bank are among the most widely known intergovernmental organizations. The fifth and final donor is that of governments.

Interestingly, the last two types of donors (intergovernmental organizations, and governments) appear to draw the most attention in existing literature (Brinkerhoff & Brinkerhoff, 2004; Krause, 2014; McEwan & Mawdsley, 2012). This is due in large part to the large amount of money these donors provide as well as their underlying political interest. When I refer to donors in this paper, I am generally referring to all five of these donor types. When I refer to a specific type of donor in my research, I will clearly state which donor type I am referring to. I have decided to do this because my research question does not specifically address a certain type of donor. Rather, I want to view donors from a broad scale as it will provide a comprehensive understanding of the general relational dynamics between donors-INGOs, and beneficiaries. Focusing on just one type of donor closes the opportunity to understand the general relational power dynamics in the donor-INGO-beneficiary system. While I think that this could be important to research, it is not this paper's goal to tackle this topic. This paper's goal is to portray the "big picture" rather than one part of the painting.

Types of Accountability: Social Accountability

The term "accountability" within the context of INGOs was born out of the term sustainability and sustainable development. These concepts did not enter regular vocabulary until the 1980s (Detraz, 2017). While a number of scholars situate sustainable development within an environmental context, this paper focuses on a social context. The definition of sustainable development can be applied to social sustainability as well.

There is no shortage of research on social sustainability and accountability. Most of the literature on these concepts center more on the concept of accountability (Bennett, 2010; Gibson, 2019; Jordan & van Tuijl, 2006; Lindenberg & Dobel, 1999; O'Leary, 2017; Scott-Villiers, Unerman & O'Dwyer, 2010). Alan Fowler was one of the first to tackle these two concepts (social sustainability and accountability) head on. His general argument is that in order for INGOs to be effective, they must "maintain the right balance between the contradictory forces, expectations, demands and processes associated with performing complex tasks in collaboration with resource-poor, powerless people in unstable and often hostile en-

vironments" (Fowler, 1997, p. xiii). Fowler's thoughts do not explicitly mention the concept of social sustainability, yet it is clear his work laid the groundwork for the concepts of social sustainability and accountability. This is because Fowler directly addresses both the relational power dynamics between INGOs and their beneficiaries and the need for effective INGO management systems in order to deal with these dynamics effectively (Fowler, 1997).

There are several types of accountability within an INGO context. For example, a hybrid term "social accountability" combines the concepts of social sustainability and accountability. This term refers to the INGO-beneficiary relationship (O'Leary, 2017; Scott-Villiers, 2002). This is similar to holistic accountability which will be explained shortly (Unerman & O'Dwyer, 2010). Social accountability seeks to balance emphasis of upward accountability (INGO accountability to donors [Unerman & O'Dwyer, 2010]) and downward accountability (INGO accountability to their beneficiaries [Unerman & O'Dwyer, 2010]). Social accountability programs within INGOs can take many forms such as participatory reviews by beneficiaries of INGO programs, complaint systems for clients to express their concerns with the INGO, and focus groups, among others (O'Leary, 2017). Though these systems may appear beneficial on paper, they often have flaws. For example, studies have shown that INGO social accountability programs have low beneficiary participation due to the beneficiary's reluctance to criticize the INGO (O'Leary, 2017). This can be because of fear of losing services that is offered by INGOs (O'Leary, 2017), or because of cultural reasons (Krause, 2014).

Additionally, there have been case studies of financially prominent INGOs which have provided valuable insight into how some of the largest INGOs view social sustainability. By interviewing various staff members from three major INGOs and focusing on their thoughts on aid dependency during complex political emergencies (CPEs), Hughes uncovers groundbreaking insights into how INGOs deal with social sustainability. For example, one of the INGO staff members interviewed by Hughes admitted that there appeared to be a conflict between donor and NGO staff interests. The staff member admitted that the NGO staff wanted to pursue sustainable development whilst donors seemed to be content with keeping beneficiaries dependent on aid so that the donors could manipulate them (Hughes, 2004). Hughes' findings also reveal important aspects of the power dynamics within INGOs which will be discussed later in this section.

Additional Types of Accountability

Social accountability, upward accountability, and downward accountability are not the only type of accountability that exists within INGOs. Some other noteworthy accountability types are "mechanisms" for NGO accountability (Unerman & O'Dwyer, 2010). These include identity accountability, and holistic accountability. The types of accountability are explained in Table 1. Whilst holistic accountability seems to be the same as social accountability, there is an important distinction between the two. Social accountability puts much more emphasis on downward accountability while holistic accountability advocates for a balance between upward and downward accountability.

Type	Description
Upward	Accountability to donors (Unerman & O'Dwyer, 2010)
Downward	Accountability to beneficiaries (Unerman & O'Dwyer, 2010)
Identity	Accountability to the INGO's right to do the right thing by delivering aid (Unerman & O'Dwyer, 2010).
Holistic	A mix between upward and downward accountability (Unerman & O'Dwyer, 2010).
Social	A mix between upward and downward accountability with seemingly more emphasis placed on downward accountability (O'Leary, 2017; Scott-Villiers, 2002)

Table 1.
Types of Accountability in a Donor-INGO-Beneficiary Context

Human Rights-Based Approaches (HRBAs)

Incorporated within the ideas of social accountability is the concept of human rights-based approaches HRBAs to aid and development. HRBAs relate to social accountability in that there is a shift of focus from aid as a charity to aid as a right (Hughes, 2014; Unerman & O'Dwyer, 2010). The mantra has been to view aid as a private charity (Hughes, 2004). Yet if aid was viewed and given as a public good, there can be much better regulation of the humanitarian aid industry (Hughes, 2004).

There appears to be a disconnect between development policy and human rights (D'Hollander et al., 2013). Just as the various types of accountability have their respective weaknesses, so do HRBAs. For example, HRBAs are very difficult to implement, especially in large NGOs. Donors of INGOs to this day still lack the ability to implement HRBAs in their policies (D'Hollander et al., 2013).

The concepts of social sustainability and accountability are best visualized with Munasinghe's "sustainomics" which includes a "triangle" of sustainable development. This triangle includes economic, social, and environmental systems (Munasinghe, 2009). Munasinghe argues that the triangle of sustainable development "can improve decision making, strengthen ownership and include poor and disadvantaged groups" which indeed falls under social sustainability and accountability (Munasinghe, 2009, p. 132). As stated earlier, this paper is interested in examining sustainability purely through a social lens. This does not mean that the other two parts of the sustainomics triangle are less important.

Literature Review

When studying the international system and the INGO world, there are a host of terms, concepts, and themes from literature that bear explaining and elaboration. It is important to note that the vast majority of these themes, concepts, and terms are interrelated and often overlap. Therefore, grouping this literature review into independent sections is quite difficult. While the following literature review is separated into sections, interconnection between the concepts should not be overlooked. The basic structure of the following literature review consists of eight sections on the themes within current literature on the topic, followed by a section on the synthesis of the current literature and then how my research will fit into the existing literature.

A Brief History of the INGO

Today, there are numerous INGOs operating all across the globe. Many of these INGOs manage funds that are greater than those of the countries they operate in (Gibson, 2019). This has not always been the case, however. How did the magnitude of INGO influence and reach get to where it is today? Several scholars attribute it to "reconstruction, decolonisation, and the collapse of the iron curtain" (Gibson, 2019, p.9; Lindenberg & Doble, 1999).

When looking back on the history of the modern INGO, a lot of scholarship points to the Red Cross as being one of the first INGOs established. However, at the time of the Red Cross' founding in 1869, the term "nongovernmental organization" was not coined. In fact, "NGO" was not coined until 1945 by the UN (Gibson, 2019). Yet, the concept of humanitarianism and humanitarian work has been around for much longer (Krause, 2014).

Since the creation of the INGO as an institution in society, neutrality has been a major debate among INGOs (Gibson, 2019; Krause, 2014). This debate is crucial to understanding INGOs because it has been a debate that continues to shape and change how INGOs operate. The debate highlight an important, wider debate on NGOs—How do INGOs choose who to serve and how do INGOs cope with not being able to serve everyone in need (Krause, 2014)? This view on the history of INGOs looks at certain historical events.

Another take on the history of the INGO focuses not on specific historical events, but on the history and evolution of INGO humanitarian practices. A main issue for this view is that of transition from short-term aid to long-term aid (Audet, 2015; Scott-Villiers, 2002). Eventually, this problem of transition led to the formation of the development "contiguum" beginning in the 1990s. This idea suggested that INGO aid strategies should be founded on the concept that relief, rehabilitation, and development approaches could be executed at the same time (Audet, 2015).

To solidify the observation that development and humanitarian ideal INGOs are relatively new to society today, it is worth learning the history of the terms "international development" and "community development". These terms only entered the INGO arena during the decolonization of Africa (Gibson, 2019). From this point onward, INGOs continue to develop their organizational structures into what Gibson classifies as "large, hierarchical organisational structures, managing substantial resources" (Gibson, 2019, p.18).

The Case Against the INGO

There is indeed a body of literature that has pushed forth the idea of abandoning the INGO altogether because of the apparent inability for INGOs to implement successful social sustainability and accountability (Bennett, 2010). According to Bennett (2010), INGOs are only accountable to their donors and therefore only focus on record keeping instead of impact. In other words, INGOs are solely accountable to where the money comes from: donors (Bennett, 2010). This argument is completed by encouraging non-NGO alternatives to development such as social businesses, patient capital investing, and microfunding to name a few (Bennett, 2010).

It has been discovered that donor interest plays a major role in the creation and implementation of relief projects of some of the major relief organizations in the world today. Even the perception of donor interest, even if a false perception, affects INGO projects to a degree (Krause,

2014). While the reasoning behind the abandonment of INGOs is valid, and there is much to learn from these viewpoints, this paper is written from the standpoint that these conclusions are overly pessimistic. This paper is more closely associated with Krause's (2014) take on this issue.

Power Dynamics

Power dynamics are also critical when understanding social sustainability and INGOs. Power dynamics among INGOs are extremely variable, delicate, and complex across political, cultural, and organizational contexts. The delicacy of power dynamics is exemplified in the often-extreme vulnerability of beneficiaries. Some scholars see these power dynamics as domination in which beneficiaries are at the complete discretion of aid agencies and their staff (Krause, 2014).

Power within INGOs is also conceptualized as the "P-word" because power is often central to INGO work, but it is not directly addressed (Eyben, 2006, p.5).[1] The "P-word" implies the inability of INGOs to address the issue of power in aid relationships. To correctly deal with power in the context of humanitarian aid, Eyben suggests INGOs should consider developing organizational cultures centered on learning and accountability (Eyben, 2006).

Donor-INGO Relationship and Power Dynamics

Donors of INGOs usually have significant power on the implementation of INGO strategies. Government donors often use their financial leverage over INGOs to push forth a certain political agenda (Jordan & van Tuijl, 2006; Krause, 2014). Similarly, corporations who are donors to INGOs use their leverage to advance a marketing agenda (Jordan & van Tuijl, 2006). In contrast to these agendas, INGOs have their own which include providing effective services to their beneficiaries.

Power dynamics are also visualized solely through a political lens. Scholars in this camp point to Trilateral Development Cooperation (TDC) for example. In TDC a country from the global North, or a nonstate actor such as the UN, partners with an "anchor" country from the global South

[1] Eyben formulated the idea of the "P" word from a statement a manager in the Department for International Development (DFID) said: "'Power is at the heart of what we do; but it is an invisible word'" (Eyben, 2006, p.5). Eyben explains a workshop she was a part of in 2001. This workshop was designed to discuss how relationships relate to power in an aid context (Eyben, 2006). It was through this workshop that Eyben discovered that many aid agencies' staff are unfamiliar with criticisms on development strategies in terms of power relationships (Eyben, 2006).

to work with a recipient country that is also in the global South. TDC is intended to improve development relationships by equalizing development power dynamics. While TDC focuses primarily on political and state actors, nonstate actors (like INGOs) can also take part in TDC (McEwan & Mawdsley, 2012).

Similarly, scholars have examined donor organizations in order to evaluate power dynamics within the INGO context. For example, USAID, the UN, and the World Bank have been studied by looking at how these intergovernmental organizations relate to NGDOs as donors (Brinkerhoff & Brinkerhoff, 2004). It has been discovered that bureaucracy and stringent reporting systems on the part of the donor ultimately hinders nongovernmental development organization (NGDO) effectiveness (Brinkerhoff & Brinkerhoff, 2004). These findings highlight that power dynamics does not only concern power through money, but also power through administrative means. Donors often have a strong say in how INGO work is monitored and evaluated.

Power Dynamics in INGO Media

Another important theme to understanding the power dynamics in INGO relationships with donors and beneficiaries is the portrayal of beneficiaries by INGO media. This is often referred to as "pornographic" representation due to its tendency to exploit the lives of beneficiaries for the sake of fundraising (Bhati, 2021). Recent studies have shown that to this day, beneficiaries are often negatively stereotyped and victimized excessively (Bhati, 2021). Issues like these are often seen as stemming from colonialism (Bhati, 2021; Krause, 2014).

This discussion of power dynamics shows that donors tend to have large influences over INGOs while beneficiaries usually do not. This leaves a gap in power dynamics between them. INGO beneficiaries are not only often stereotyped by INGOs (Bhatai, 2021), but also are often used in competition with other beneficiaries to receive funding for certain INGO projects (Krause, 2014). This is precisely where healthy accountability and socially sustainable practices come into play; they exist to bridge this gap. This realization has led to the recommendation for INGOs to take a more bottom-up approach to development (Lewis et al., 2021).

Systems Thinking

A systems thinking approach was not developed as a part of the fields of study that this paper draws from. Yet, this type of thinking can be applied to a variety of disciplines (Arnold & Wade, 2015). Originally devised in

1987 by Barry Richmond, "systems thinking" is, "literally a system of thinking about systems" (Arnold & Wade, 2015, p.670). Arnold and Wade (2015) also provide a "systemigram" which pictures, in diagram form, the relation between elements and interconnections of a system (Arnold, 2015, p.675-676). In this paper, I pose that by using systems thinking to understand INGOs and their relationships with donors and beneficiaries, it will be much easier to situate the concept of social sustainability into this discussion. Systems thinking will be used in the conclusion section.

Synthesis of Existing Literature

Given that INGOs, especially aid and development ones, are relatively new in societies today, the academic research around these entities is ever evolving and growing. This also means that there are several noticeable gaps in existing literature when it comes to the donor-INGO-beneficiary relationship and social sustainability.

First, there is not much literature specifically relating social sustainability to INGOs. Instead, terms like "accountability", "social accountability", and "sustainable development" are all discussed within the INGO context. These terms all have similarities to the concept of social sustainability in their respective ways. However, the discussion of social sustainability and INGOs is rare in existing literature.

Second, there is not much literature regarding the changes in how INGOs approach the transition from short-term aid to long-term development. At what point should INGOs stop providing short-term relief and begin long-term development? What does this transition look like? Why are so many relief INGOs expanding into development work? Is there even such thing as a relief organization given that many INGOs are expanding into long-term development work? Questions like these seem to expose a gap in existing literature. There are many bodies of work that seek to focus solely on humanitarian relief organizations. However, these same bodies of work do not attempt to thoroughly explain the fact that so many of the INGOs they are studying would also be considered by many to be development INGOs. The terms "aid" and "development" are often blended together and used interchangeably in existing literature, leaving room for misinterpretation and obscurity.

Research methods in current literature also reveal a gap in existing literature. Much of the research done on the topics of this paper are in the form of case studies. These case studies vary in focus. Some case studies are on INGOs as a whole while others are on specific INGO programs. Interviews as a research method are also very common in existing literature.

While each of these methods are extremely valuable, there appears to be a relatively weak variety of research methods within existing scholarship.

Situating My Research

I hope to fill in some of these gaps through my research and findings. I hope to do this first by specifically contextualizing the term "social sustainability" within the existing body of scholarship. As noted above, there is not a lot of scholarship exactly referring to social sustainability. I will argue that social sustainability is an important term to discuss when looking at INGOs and it most definitely belongs in the current body of research.

Second, my research uses a relatively unique methodology in relation to the research that has already been done on the topics of INGOs. I will be looking at how INGOs themselves discuss their relationships with their beneficiaries and donors, and how they perceive the topic of social sustainability given these relationships. To do this, I have completed a qualitative content analysis of INGO websites. This research method diverges from what I consider the norm under the topic of INGOs. I have not done a case study of specific INGO programs. Nor have I conducted interviews with INGO staff. Instead, I am focusing on the media produced by INGOs. By focusing on this, I hope not only to add to the variety of research on this topic, but also to fill in and enhance the current body of literature.

Methodology

To answer my research question, I have chosen to do a qualitative content analysis of eight INGO websites. In this section I will provide a thorough explanation of this method, including how I chose the eight INGOs in my study. The analysis section will explain the specific findings of my research. To understand and situate my findings, I will use systems thinking which will be applied to INGOs and their relationships with donors and beneficiaries. To accomplish this, the donor-INGO-beneficiary relationship will be evoked as a system in and of itself. I argue that by following Arnold & Wade's (2015) systems thinking approach, the complexities of INGOs and their relationship with donors and beneficiaries can more easily be understood. Likewise, understanding INGOs through systems thinking also helps better understand their conceptualizations of social sustainability. All of this will also help discover potential change, if any, for INGOs regarding social sustainability.

In particular, I will focus on the third element of Arnold & Wade's (2015) "systemigram" which is, "understanding system structure" (Ar-

nold & Wade, 2015, p.676). This understanding improves "the capability of identifying and understanding systems, predicting their behaviors, and devising modifications to them in order to produce desired effects" (Arnold, & Wade, 2015, p.676). For the purpose of answering the research question, this quote will be split into three sections: (1) "identifying and understanding systems", (2) "predicting their behaviors", and (3) "devising modifications". The conclusion to this research will be modeled after these three sections.

Research Methods

My qualitative content analysis, its framework, and its structure are adopted primarily from the work of White & Marsh (2006). While quantitative content analyses focus on testing a particular hypothesis, qualitative analyses include "foreshadowing questions" which are "open questions that guide the research and influence the data that are gathered" (White & Marsh, 2006, p.34). As will be explained shortly, my qualitative content analysis is precisely based on such questions. Qualitative content analyses are fluid and can be shifted and modified as the researcher observes the patterns in the text being studied. Moreover, qualitative analyses are focused on understanding the big picture and general patterns as opposed to specific hypotheses (White & Marsh, 2006). The use of numbers and percentages can be used in qualitative content analyses, yet specific, detailed statistics are not usually included (White & Marsh, 2006). Also, data collection and presentation commonly come in the form of blocks of texts that answer the researcher's guiding questions (White & Marsh, 2006). These characteristics of qualitative analyses will be clear throughout my methodology and analysis section.

I chose to do a qualitative content analysis because of the flexibility that is given to the researcher. This flexibility allowed me to adjust my research according to what I was learning from the INGO text I was studying. In other words, this method allowed me to focus on context-specific questions that many other research methods do not have the flexibility to address. Furthermore, the terms and concepts of study in this paper (social sustainability, accountability, etc.), as will be seen shortly, are widely interpreted. Therefore, studying these terms through fairly rigid quantitative analysis methods would prove quite difficult. The final reason I chose a qualitative approach to this paper is because I wanted the ultimate goal of this paper's research to be inductive not positivistic. This key difference between quantitative and qualitative content analyses is outlined in the work of White and Marsh (2006).

There are certainly both positives and negatives to doing a qualitative content analysis. One positive was mentioned above: the flexibility the researcher has in answering the particular research question. This flexibility allows researchers to modify and shift their research according to patterns they identify along the way which is absent in many other qualitative and quantitative analyses (White & Marsh, 2006). This ultimately allows researchers to fill in gaps in existing literature and research which is something that other research methods cannot accomplish. While the flexibility of qualitative content analysis is a major positive, it can also be a negative. A qualitative content analysis is completely dependent on the bias and personal views of the researcher which can consequently result in bias in the actual research findings.

INGO Selection

To perform a qualitative content analysis of INGO websites, it must first be decided which websites to choose from. This is an especially daunting task given the diverse and vast INGO world. There are two criteria I used to determine which INGOs I should study. The first criterion is the requirement that the INGO must be both be headquartered in what many scholars consider the global North and have an office in the USA.

I chose this as a criterion because INGOs headquartered in these countries are significantly challenged with social sustainability as seen in the examples in my introduction. It is these INGOs that have grabbed headlines surrounding their scandals and controversies relating to social sustainability. It is these INGOs that also struggle with relational power dynamics because of the wealth of the countries they were founded in. Being headquartered in the global North often implies that these INGOs are dealing with a significant amount of donor funds. Many of these INGOs deal with hundreds of millions of dollars of funds annually. Additionally, much research has been done on these INGOs given their financial prominence. I want to complement this existing body of research. My choice to focus mainly on Northern INGOs is also reflected in the literature I have cited in this project. Most of the literature cited refers to Northern INGOs.

There is also a practical reason why I chose this criterion. At the time of my research, I did not have the time nor resources to examine what many scholars consider grassroots NGOs or NGOs that were founded and headquartered in the global South. It is simply too difficult to study every type of INGO as a part of one research project. Language barriers, and difficulty in locating their websites on American servers are just a few of the challenges in studying these INGOs. I am not attempting to ignore the sig-

nificance of Southern based INGOs. In fact, studying grassroots and global South INGOs is extremely important. I simply did not have the resources to include these organizations in my research. Therefore, when I refer to INGOs at any point in this paper, I am referring to Northern-based INGOs.

Another criterion I used to decide which INGOs to study is the requirement that the INGO must have a mission statement and provide programs that are specifically aimed at humanitarian aid and development. Therefore, when I use the term INGOs throughout my paper, I am referring to INGOS which are focused on humanitarian aid and development only.

I recognize that this criterion can result in implicit bias on the researcher's part because of the debate over what is considered aid and what is considered development. I view humanitarian aid as that which is used for short-term goals, while development is more focused on long-term goals. This view is consistent in various bodies of the literature surrounding INGOs (Audet, 2015; Krause, 2014). Disaster relief and crisis response are common humanitarian aid programs. Education, and economic advancement are a few examples of long-term development programs.

I will use the INGO Medecins Sans Frontieres (Doctors Without Borders) as an example to clarify this criterion: MSF has headquarters in the global north (Switzerland) and has an office in the USA. The mission of MSF is to bring "medical care to people affected by conflict, disasters, epidemics, and social exclusion" (Medecins Sans Frontiers, n.d.). For my research, MSF meets the first of my criteria (headquartered in the global North and has an office in the USA). However, as seen in its mission statement, MSF is only focused on humanitarian aid (in the form of medical care) and not development. Therefore, MSF does not fit my second criterion (to be involved in humanitarian aid and development work) and consequently, I did not include MSF in the scope of my research. These two criteria have allowed me to focus specifically on a type of INGO that, as I have argued in this paper, are faced with significant challenges concerning social sustainability and relational power dynamics.

Nevertheless, these two criteria narrow down the list of possible INGOs only to an extent. There are still countless INGOs that meet these two criteria. When deciding to choose which INGOs to study I initially did a general web search on the largest INGOs in the world. There are several different websites that attempt to list the biggest INGOs. I quickly learned that it is extremely difficult to create a ranking of the largest INGOs in the world. Each source I reviewed had variations in their lists. Moreover, I did not have the time nor resources to create my own comprehensive list.

After evaluating various lists from different sources, I eventually came across an article from Human Rights Careers (HRC). As implied in their name, HRC has the goal of providing career resources for human rights students and professionals (Human Rights Careers, n.d.). Additionally, HRC appears to have a good reputation in the human rights field. As stated on their website, some of their most prominent clients are the United Nations, various NGOs, as well as various universities across the globe (Human Rights Careers, n.d.). Because of these reasons, I chose to use HRCs article titled The Fifteen Biggest NGOs in the World to further narrow down my choices for research. I noticed that in this list were many INGOs who have already been the center of discussion in some of the literature I cited. This confirmed my decision to use this article. From the fifteen INGOs listed in the HRC article, I was able to conclude that eight of them fit my two criteria. So, I chose to study those eight INGOs. In order of their ranking on the HRC article, they are: Save the Children (rank 1), Oxfam International (rank 2), World Vision (WV; rank 5), International Rescue Committee (IRC; rank 6), Catholic Relief Services (CRS; rank 7), CARE International (CARE; rank 9), ActionAid International (rank 11), and Plan International (rank 15).

Once more, I recognize that what are considered the largest INGOs in the world are widely debated and vary by source. When deciding which INGOs to research, I recognized this challenge, and I recognized that my choices of which INGOs to research can be debated and contested. But, as I have just explained, I argue that my process to choose these INGOs was carefully thought out and executed with stringent and logical criteria. For reference, the INGOs that did not meet my criteria but were listed on the HRC article, in order of HRC's ranking, are MSF (rank 3), BRAC (rank 4), Danish Refugee Council (rank 8), Amnesty International (rank 10), Direct Relief (rank 12), Action Against Hunger (rank 13), and Anti-Slavery International (rank 14).

Research Process

To perform a qualitative content analysis of these eight INGO websites I centered my research process on pre-written questions as explained earlier in the work of White & Marsh (2006). I wrote a series of questions that I believe would best answer my research question. Some of the questions I used include: What does the INGO say about social sustainability on their front page? Do INGOs appear to put more focus on donors or beneficiaries? How do INGOs discuss social sustainability (for example, what medium do they use)? What key terms do INGOs use to discuss social sus-

tainability? Is it easy to find information about the INGO's social sustainability practices? When talking about social sustainability, how specific do INGOs get? How much text do INGOs dedicate to social sustainability? Are there any changes over time to how INGOs discuss sustainability?

I came up with these questions after first doing a general and brief examination of the INGO websites to get an idea of some potential patterns across them. This allowed me to discover some context for my research questions. Once again, these questions are dependent on my own research interests and bias. However, I argue these questions are comprehensive because they thoroughly address the terms and concepts that are the focus of this paper (for example: social sustainability, INGO-donor-beneficiary relationships, power dynamics, etc.). These questions do not merely address the concepts outlined in this paper, but they address them in a multi-faceted way. The questions address the topic of social sustainability through medium (type of text), the ease of access to social sustainability text, and the specificity of such text. The variety of approaches these research questions take to analyze social sustainability on INGO websites allows for a broad and comprehensive understanding of social sustainability in the INGO context.

I first did a basic exploration of the front page of each INGO website. I did this because I operated on the assumption that the front page is basically the first impression the INGOs provide to the world. As a result, I believe the INGOs would put what matters to them most on the front page so that viewers can understand the INGO in a matter of minutes. I also found and analyzed each of the INGOs' mission statements to ensure that they met the criteria outlined earlier. It is important to note that the INGOs in my research are so large that many of them have multiple websites. When possible, I used the INGO's international website. However, when this was not possible, I used the INGO's USA website for reasons of accessibility.

After exploring the front page, I explored the rest of the website in search of key terms that fit under the umbrella of social sustainability. I quickly learned that many of these INGOs mostly use annual reports, accountability reports, financial reports, or strategic plans to discuss social sustainability. These texts make up the bulk of my analysis. I mostly focused on the most recent of these texts whenever possible. This method I used is consistent with White and Marsh's (2006) explanation that in qualitative content analyses the researcher may shift the focus of the research based off of observed patterns. Within these texts, I used power searching tools to find key terms all while answering the questions I listed

above. I analyzed the data (the answers to my questions) and tried to find common themes, similarities, and differences across the INGOs I studied. Once again, this technique is explained by White & Marsh (2006) where researchers conducting qualitative content analyses seek the big picture and general patterns. For context, I discovered that annual reports are generally aimed at explaining the INGOs work and progress in the past calendar year. Accountability reports are documents outlining the measures the INGO has taken to self-regulate its operations. As their name suggests, financial reports in the context of INGOs are documents outlining the donor money the INGO has received and how that money has been spent.

Analysis

As noted above, my qualitative content analysis is based on eight questions regarding each of the INGO's websites. I have looked at how social sustainability is conveyed on these websites. The eight questions are organized into three different groups. Each group has a certain goal in answering the overall research question of this paper. I classify each group of questions as a theme because each group is centered on a specific aspect of the research question. Henceforth, I will use an abbreviation for social sustainability: SS.

The first theme includes three questions related to the front page of the INGO's website. These questions are intended to examine how easy it is to find information on SS. I interpret the ease of finding this material as the extent the INGO places importance on displaying their SS practices. In other words, if the INGO believes that it is important for viewers to know how they ensure SS practices, then the INGO will make such text easily accessible via the website. The three questions for this theme are: (1) Can SS text be accessed from the front page? (2) Is the front page focused more on donors or beneficiaries? (3) How easy it is to find SS material?

The second theme also includes three questions and is the most specific to the research question at hand. While the first theme is a general analysis of the front page, this second theme is focused solely on the INGOs' SS texts and is designed to understand how the INGOs discuss, interpret, and understand SS. These questions include: (1) What are the ways/mediums in which SS is discussed? (2) What are some of the key terms that the INGO uses to discuss SS? (3) How specific and detailed does the INGO get when discussing SS?

The third and final theme includes two questions. This theme has a broader context and seeks to understand the overall trends in the INGOs' conceptualizations of SS. The two questions in this theme are: (1) What

is the amount of text that is dedicated to SS? (2) What are any noticeable changes to how SS is conceptualized on the INGOs' websites? The bulk of my findings regarding this third theme will be discussed in the conclusion section as they refer to broader observations.

Given the three themes of my research findings, this analysis section is structured accordingly with the name of the three themes as subheadings. Under each theme's subheading, I will first provide a table summarizing the notes I took during my research process. I will then explain the main findings for that respective theme. My main findings, which are in the form of phrases and statements, will be italicized to help readers follow along more easily and to help in their understanding of what is being discussed.

	SS on Front Page?	Focus on beneficiaries or donors?
Save The Children	-Intl. site: hyperlinks to "accountability" and "safeguarding" at bottom of the page -US site: mostly info on donating.	-Intl. site: broader; section on donating, news briefs, newsletter, etc. -US site: heavy focus on donating (donation graphic)
Oxfam	-Intl. site: Hyperlink at bottom of page called "our finances and accountability" -US site: tab on safeguarding.	-Intl. Site: Not much info on the front page about donating or donors -US site: more donation options
World Vision	-Intl. Site: Hyperlink at the bottom of page called "accountability" -US site: only offers information on financial accountability	-Intl. Site: seems to recognize the power dynamics ("Transformed relationships" driver of sustainability) -US site: Heavy emphasis on donating.
IRC	-US and Intl. Site are the same. -No accountability tab or hyperlink.	-Shows a celebrity donor on the front page. -In "how we use evidence" tab, IRC directly mentions how donors want accurate reporting systems.
CRS	-Hyperlink/tab for "safeguarding". -Hyperlink to "research and publications" where SS material can be found.	-Main site has multiple areas to give.
CARE	-Accountability and transparency hyperlink at the bottom of the front page.	-Several sections of the website are dedicated to donating.
Plan International	-Accountability hyperlink at the bottom of the front page.	-Much of the front-page content is dedicated to beneficiaries. -Less emphasis on donors
ActionAid	-Intl. Site: Safeguarding hyperlink at the bottom of the page. -US site: SS not evident and accessible from front page. May be due in part because the US branch is concerned with advocacy.	-Intl. Site: Similar to US site. -US site: Lots of options to donate on the front page.

Table 2.
Theme 1 (Accessibility of SS Text) Notes

Under this first theme, a main finding is that in most cases, some form social sustainability content is accessible from the front page. The most common way to access this text from the front page is through a hyperlink

at the bottom of the page. Many of the hyperlinks are titled "accountability", "safeguarding", "transparency", or even a combination of these terms. Some websites have multiple hyperlinks at the bottom of the page that guide website users to various SS material. Once these hyperlinks are clicked, the viewer is taken to a page where SS material can be found. The only exception to this is the IRC's website. There is no hyperlink to SS material on IRC's website. Figures 1 and 2 (depicted below) are examples of hyperlinks from different INGO websites. Figure 1 depicts the bottom of Save the Children's international website. As seen in this picture on the right-hand side, there are two hyperlinks to SS material— "Accountability" and "Safeguarding") (Save the Children, n.d.). Figure 2 depicts the bottom of CARE's website. The third column in this figure has a hyperlink titled "Accountability and Transparency" (CARE, n.d.).

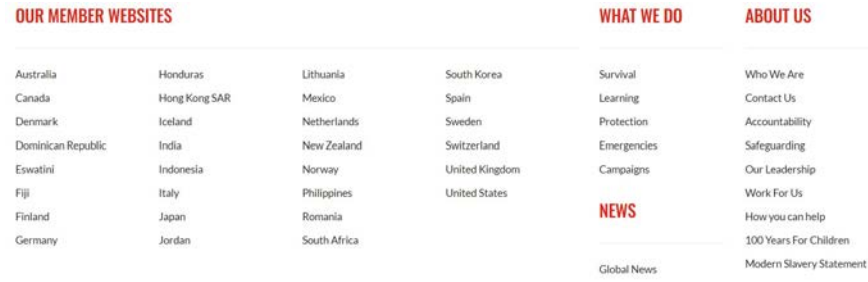


Figure 1.
The Bottom of Save the Children's Front Page

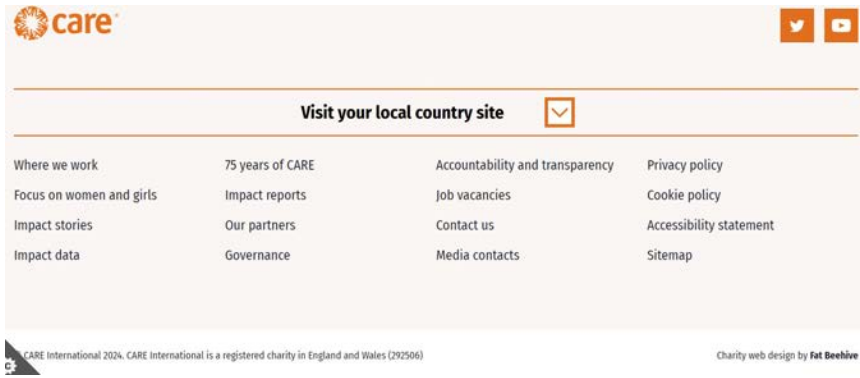


Figure 2.
The Bottom of CARE's Front Page

In cases where the US website and international website are different, as in the examples of ActionAid, World Vision, Oxfam, and Save the Children, the US sites do not have hyperlinks to SS material on the front page except for Oxfam. In many cases, the bottom of the US websites includes a graphic on financial reporting. This is interesting because the US sites appear to focus less on SS and more on financial reporting and transparency. For example, Figures 3 and 4 is a comparison between the bottom of World Vision's US site (Figure 3) and their international site (Figure 4). As seen in this comparison, the US site has a graphic showing the flow of money in the organization. However, the international site does not have such a graphic (World Vision, n.d.; World Vision International, n.d.).

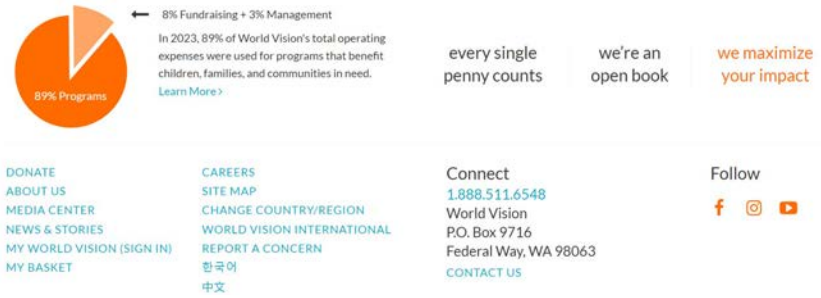


Figure 3.
The Bottom of World Vision's US Front Page

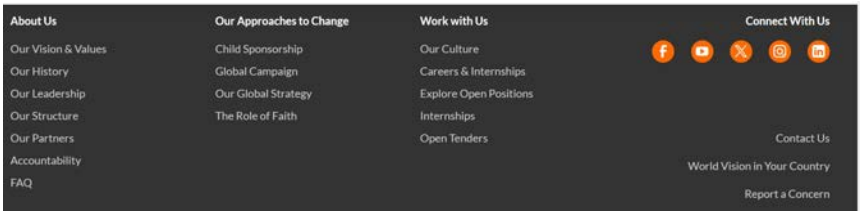


Figure 4.
The Bottom of World Vision's International Front Page

There are a host of potential reasons why the US sites tend to focus less on SS and more on financial reporting, and I will explain a few of them. First, it could be simply because the US sites are satellite websites that are not the center for such information. It could also be that the INGOs view their international websites as the ones responsible to hold all the detailed reports of SS, not the US sites. Indeed, this seems to be the

case in several INGOs in this study. Save the Children, World Vision, and Oxfam are a few of the INGOs that have the bulk of their SS text located on their respective international sites.

Third, it could be because the US offices of these INGOs have slightly different goals, aims, and structures. This seems to be true in the case of ActionAid. ActionAid's US website is more centered on advocacy work instead of on the ground humanitarian and development work. On their "Who We Are" tab, ActionAid USA states that, "We influence U.S. policy and international institutions like the United Nations and elevate marginalized voices in the halls of power" (About ActionAid, USA, n.d.). Since ActionAid USA appears to be a primarily advocative organization and not really doing "groundwork" this INGO may not see a need to have information on social sustainability on their US website.

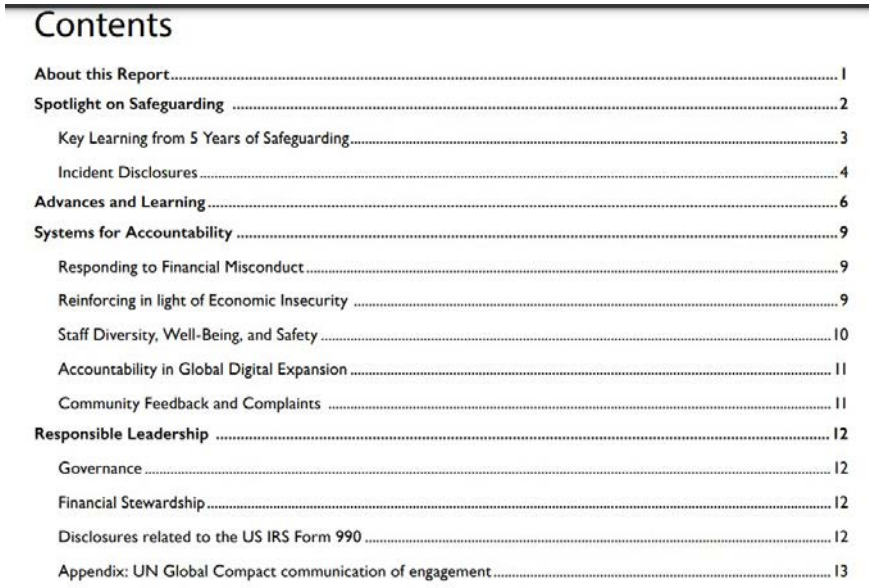
As expected, the majority of the INGO websites heavily emphasize donor participation. This was expected because it is safe to assume that these INGO websites will mostly be visited by donors and potential donors. On many of the websites, there are multiple tabs, or hyperlinks where the viewer has the ability to directly donate money via the website. Interestingly, Plan International's website is an anomaly. While the other INGOs have multiple tabs and hyperlinks for donations throughout the front page, Plan International's website only has one at the top of the page. The rest of the website is dedicated to various news, analysis, and other sources describing the issues that Plan International seeks to resolve (Plan International, n.d.).

Though the websites are inherently designed for donors, many of the front pages offer stories from beneficiaries, or other texts dedicated to explaining the humanitarian issues that beneficiaries face. However, it is difficult to discover the motive behind these texts. Are they there for donation and marketing purposes (to grab donors' attention), a view held by Bhati (2021), or are they there to empower beneficiaries by sharing their stories? This question is one to keep in mind when analyzing INGO websites and their texts regarding beneficiaries. These uncertainties are consistent with the findings of Bhati (2021) and Krause (2014) who view INGOs as engaging in a market and a competition for donor funds. INGO websites tend to clearly cater towards donors. In fact, INGOs tend to sacrifice the false stereotyping of beneficiaries in order to grab donor attention (Bhati, 2021). Moreover, the fact that the websites in this study consistently present text on beneficiaries could be viewed as the process of competition among INGOs for donors in which beneficiaries are used as a means to an end (Bhati, 2021; Krause 2014).

	Ways SS is discussed?	Key terms used to discuss SS?
Save The Children	-Accountability report -Accountability web page with hyperlinks	-Accountability, safeguarding, transparent
Oxfam	-Program case studies (specific examples of SS in specific projects)	-Accountability, human rights, learning, integrity
World Vision	-Accountability report -Directly discussed SS. -5 "drivers of sustainability"	-Empower, accountability, transparency, transformational development, sustainability
International Rescue Committee (IRC)	-Mostly through an empowerment and advocacy lens. -No apparent accountability report; relatively limited SS text	-Empowerment, social accountability
Catholic Relief Services (CRS)	-Mainly through the idea of human dignity and respect for basic human rights of beneficiaries.	-Human dignity, respect
CARE International	-Accountability reports, 2030 vision document, briefs, and other reports.	-Impact, accountability, social sustainability
Plan International	-Accountability report (published by 3rd party organization) -Pages with hyperlinks to resources, safeguarding annual report	-Rights-based, accountability, safeguarding
ActionAid	-ALPS publications, strategic vision, annual report.	-External accountability, internal accountability, human-rights based approaches

Table 3.
Theme 2 (Conceptualizations of SS) Notes

All eight INGOs in this study have separate publications on SS that are available from the website. By "separate publications" I mean text that is in the form of a document, not a webpage. A webpage implies more general text that is lacking in specificity. A separate document, like a report for example, implies a much more detailed text. Of these eight INGOs, four of them (Save the Children, WV, Plan International, and CARE) have published "accountability reports" which are rich with SS text. These reports are inward looking, meaning that they seek to examine and measure the INGO's own SS practices. These reports usually include data on safeguarding (the systems put in place to protect the INGO's staff and beneficiaries from abuse), fraud reporting, and various frameworks or approaches to aid and development. Figure 5 depicts the table of contents for World Vision's 2022 accountability report (World Vision International, 2023). The table of contents provides insightful information as to how this INGO, and many others in this study, conceptualize and present SS.



Contents

Figure 5.
World Vision's 2022 Accountability Report Table of Contents

Other common mediums used to discuss SS are strategy reports or vision statements as in the example of CRS, ActionAid, CARE, and Save the Children. These publications are more general in nature and do not "look inward". Instead, they are generally made up of statements or goals that show the INGO's dedication to SS. These publications are forward-looking, meaning that these publications describe the INGOs stance on SS and how they aim to execute their stance in the coming years. Figure 6 is an example of a strategy and goal-oriented document that describes an aspect of SS. The document was published by CRS. As seen in the second row, an aspect of SS is being discussed ("community member participation"). This document is more general in nature. It is forward looking rather than inward looking. Though SS is being discussed at some points, it is not being discussed within the context of how CRS is carrying out and executing SS practices (Catholic Relief Services, 2020). Instead, CRS is merely explaining their commitment to SS practices. Many INGOs in this study have similar documents and publications.

Figure 6.
Segment of CRS' SDG Document

Two of the INGOs in this study, Oxfam and the IRC, conceptualize SS slightly differently.[2] This is primarily because of the way their websites are structured. Oxfam and IRC are very project oriented, meaning that they focus on examining the effectiveness and impact of specific projects. For example, Oxfam has a database of project assessments as opposed to a single, organization-wide report on accountability (Accountability and Transparency, n.d.). When Oxfam participates in "inward looking" it is in the form of an "integrity report" which is primarily focused on safeguarding and fraud reporting (Oxfam, 2022).

Several of the INGOs that do not have their own accountability reports have published resources for SS as in the examples of CRS, Action Aid, and the IRC. These three organizations are interesting in that it was difficult to find specific texts that were inward looking. However, these INGOs have published various resources relating to SS. These resources come in many different forms. Most are internally reviewed. A common publication from these organizations is an explanation of a certain organizational system as in the example of ActionAid's Accountability, Learning, and Planning System (ALPS), as well as CRS' "Respect for Dignity in Daily Life Index" (Hembling et al., 2024). Sources like these are aimed at explaining the systems the INGO has in place to ensure SS rather than reporting on the impact of these respective systems.

[2] It is important to note that in many of these findings, the IRC is an anomaly in that it was quite difficult to find SS text. The IRC neither has an accountability report, nor is there a hyperlink to SS text, nor is the term "accountability" mentioned in its annual report (International Rescue Committee, 2022). The IRC does indeed have SS text but many of the publications I found were from 2015 at the latest and are not inward looking.

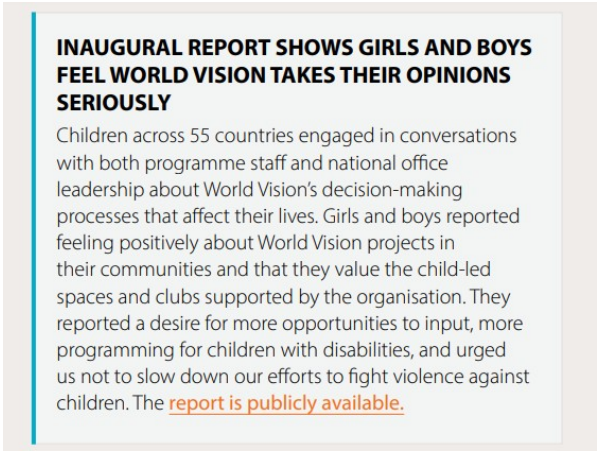
From the theme two findings explained thus far, I have come across two additional observations. The first being that the specificity on SS concepts varies greatly between the INGOs in this study. Given the variety in ways SS is discussed, it is understandable that the specificity in these discussions vary. On each of the eight INGOs websites, it was quite difficult to find empirical data to support their claims of SS. Many organizations will describe a program that is socially sustainable, or a framework for humanitarian aid that is socially sustainable, or even just talk about the ways in which they are socially sustainable. However, there appears to be a gap between the claims these INGOs are making and the evidence to support their claims.

Of the INGOs in this case study, three stood out as the most detailed in their discussions on SS. I determined these three as the most thorough in their discussions on SS because of two reasons. First, these organizations not only have large amounts of texts dedicated to SS (publications, web pages, reports, etc.), but also these texts are widely accessible via the website. Second, these organizations were some of the few in this case study that have used some pieces of empirical data to prove their SS. The three INGOs are WV, CARE, and Save the Children. By looking into these three INGOs and how they report SS practices, I was able to gain valuable insight into the ways in which SS is empirically presented.

Before elaborating on this point, I will address the fact that the majority of these organizations have empirical data on safeguarding. Therefore, the claim that all the INGOs (not just the three I chose) in this study have publicly available SS data on their websites can be made. However, to keep with the umbrella analogy, safeguarding is just one term under the umbrella of SS that characterizes one aspect of SS. I argue that the three organizations I listed above not only had data on safeguarding (which appears to be the norm), but they have empirical data on other aspects of SS too.

To better visualize the empirical data from these three organizations, I have included three figures which are images from various texts on their respective websites. The following three examples are intended to highlight how some of the INGOs in this study empirically present their SS practices. Figure 7 is an excerpt from World Vision's 2022 Accountability Report. The excerpt shows a section describing a report WV conducted. The report was aimed at gaining insight into how some of WV's beneficiaries (children in this case) view the INGO. This text describes SS because it directly addresses the INGO's impact on the communities it serves. Beneficiaries themselves had the opportunity to voice their critiques of WV's

presence in the community. Some of the findings are explained in the paragraph shown. Though there are no true "hard numbers" per se, there is an option for readers to access the actual data of the report via hyperlink (World Vision International, 2023, p.22).



INAUGURAL REPORT SHOWS GIRLS AND BOYS FEEL WORLD VISION TAKES THEIR OPINIONS SERIOUSLY

Children across 55 countries engaged in conversations with both programme staff and national office leadership about World Vision's decision-making processes that affect their lives. Girls and boys reported feeling positively about World Vision projects in their communities and that they value the child-led spaces and clubs supported by the organisation. They reported a desire for more opportunities to input, more programming for children with disabilities, and urged us not to slow down our efforts to fight violence against children. The report is publicly available.

Figure 7.
Segment from World Vision's 2022 Accountability Report

The next image, depicted in Figure 8, is an excerpt from Save the Children's 2022 Global Accountability report. In this section of the report, the INGO is describing how it is ensuring that children (Save the Children's beneficiaries) are being incorporated into INGO decision-making processes. This is seen as socially sustainable because the INGO is attempting to shift more power to beneficiaries. Similar to the example of WV, this report is directly addressing Save the Children's handling of the relational power dynamics between the INGO and the beneficiary. In the third paragraph of this text, specific empirical data is provided. The data is in percentage form and shows the percentage of Save the Children programs that are adopting "children's participation mechanisms" (Save the Children, 2023, p.16).

ENGAGING CHILDREN IN OUR OWN GOVERNANCE

To truly shift power to children, of course we must enable children to influence our own organisational strategic planning and decision-making.

Save the Children International established an Interim Global Children's Advisory Body in 2022 to empower child representatives to get involved in decision-making at a global level and to discuss issues directly with our CEO. This group has transitioned during 2023 to a Global Children's Advisory Body, with all members elected through advisory bodies at country and regional level.

In 2022, 36% of Save the Children Association entities throughout our Movement had established children's participation mechanisms. We aim to have children advising at senior leadership level in 75% of the countries in which we work by the end of 2024.

Figure 8.
Segment of Save the Children's 2022 Global Accountability Report

Interestingly, I was able to find an almost identical piece of data in CARE's accountability report. In their 2021 accountability report, CARE describes their "Feedback Accountability Mechanisms" (FAM). FAM, in a way, is similar to ActionAid's ALPS: it is a system CARE has in place to maintain SS practices. CARE provides the statistic that from 2019 to 2021, the percentage of CARE programs that have incorporated the FAM framework rose from 60% in 2019 to 72% in 2021 (CARE International Secretariat, 2021, 27). Though no generalization can be made, it is interesting that two INGOs in this study provided nearly identical pieces of data to support their respective SS practices.

Figure 9 is the third and final example of empirical data supporting SS practices. Figure 9 is an excerpt from CARE International's Sustainable Development Goals (SDG) Impact Report. This report measures CARE's impact according to the United Nation's SDGs. The sixteenth SDG, which is "peace, justice, and strong institutions" relates to SS mainly through the "strong intuitions" section of the phrase. This SDG is indeed included in CARE's report. The text shows one portion of CARE's report on the sixteenth SDG. In this text, CARE is not only providing data pertaining to SS practices, but also hyperlinks to other resources regarding the data (CARE International, n.d., p.8). As seen, the data comes in the form of numbers as well as a few percentages.

CARE contributed to an increase of over 460,000 people participating in formal or informal decision-making spaces, in in 34 countries. Ghana's Strengthening Accountability Mechanisms (GSAM) project enabled 671,880 people to participate in formal or informal decision-making spaces (42% women), an increase of 93,440 people. The Implementation of Social Accountability Framework (ISAF) project in Cambodia enabled 22,392 people to participate in formal or informal decision-making spaces (58% women), with 294 people in new leadership positions (53% women).

Figure 9.
Segment From CARE's SDG Impact Report

These are just some examples of how data on SS is reported by the INGOs in this case study. In the same sources that these excerpts came from, there are other examples of how SS data is conveyed. For example, CARE included some numbers on the increase in beneficiary involvement in decision making since 2014 (CARE International, n.d., p.8), WV provided percentages and percentage change in programs that involved community-member assessments (World Vision International, 2022, p.7). Save the Children provided numbers on the amount of funds being transferred to local partners (Save the Children, 2023, p.13).

The aforementioned figures also bring to light two brief realizations. First, INGOs in this case study appear to place at least some importance on supporting SS claims with empirical data. Second, the empirical data found from the websites in this study are sparse. This could be because of the reality that measuring SS is a daunting task. As said multiple times in this paper, SS is a very broad term with so many moving parts, making it difficult to measure SS.

The term "accountability" is by far the most common term used to conceptualize social sustainability. This finding is in concordance with existing literature that conceptualizes SS through the concept of accountability (Bennett, 2010; Gibson, 2019; Jordan & van Tuijl, 2006; Lindenberg & Dobel, 1999; O'Leary, 2017; Scott-Villiers, 2002; Unerman & O'Dwyer, 2010). All eight of the INGOs in this study have some sort of hyperlink, text title, or tab that mentions the term "accountability". Interestingly, I was able to find two instances in which the term social sustainability is directly addressed. The first instance is CARE International's 2030 vision. In this document, CARE writes, "We work to influence changes that are environmentally, socially and institutionally sustainable" (CARE International, 2021, p.5). The second instance is World Vision's web page on sustainability. On this page, WV provides their own definition of sustainability: "Sustainability refers to the ability of local communities, partners and other stakeholders to continue to maintain and improve child well-being

after the end of WV's programme interventions. This definition is rooted in the recognition that our programmes operate for a limited duration and that World Vision's contribution to a community's journey will always be temporary." (Our Approach to Sustainable Change-How do we Sustain Improvements in Child Well-Being, n.d.). Note that social sustainability is not mentioned in this excerpt, but WV's definition is almost identical to the definition of SS that this paper has laid out. This shows that WV conceptualizes the even broader term "sustainability" within a social context. WV's definition of sustainability emphasizes the underlying issue INGOs face of delivering and transitioning from short term aid to long-term development (Audet, 2015; Hughes, 2004).

All three of these examples are also similar in that they have the intention of empowering and improving the relational power dynamics between the INGO and the beneficiary. The programs outlined in these figures reveal the many similarities between two significant INGO social sustainability programs that have been the subject of analysis in existing literature—ActionAid's ALPS and Farmer Field Schools (FFS) in Kenya (Duveskog et al., 2011; Scott-Villiers, 2002).

ActionAid was one of the first to put social accountability to practice through their ALPS. To align with social accountability practices, ALPS essentially shifted ActionAid's focus on upward accountability to downward accountability. The core idea of ALPS was that poor communities are the primary stakeholders for NGO services and that services should be provided with a stakeholder first mindset. ALPS essentially eliminated the process of long forms and reports and now focuses on interviews of the communities being served. Instead of long reports, ALPS focused on learning and instead of bureaucratic procedures, ALPS prioritized principles (Scott-Villiers, 2002).

Despite these seemingly beneficial aspects of social sustainability that ALPS brings to the table, the program has potential vulnerabilities and flaws. For example, because ALPS is not based on rules, the program is open to misinterpretation, manipulation and corruption. This in turn can result in even more bureaucratic procedures than before. Nevertheless, it can be concluded that ALPS has the potential to have positive effects on INGOs (Scott-Villiers, 2002). These conclusions are important to keep in mind when analyzing the INGOs in this paper's study.

Another similar program that has been the subject of analysis in existing literature is that of Farmer Field Schools (FFS) in Kenya (Duveskog et al., 2011). The FFS programs have the intention of empowering farmers by providing a platform for them to share their experiences and knowl-

edge of farming with other farmers (Duveskog et al., 2011). Researchers discovered that FFS programs, which many development agencies are putting into practice, have proven to bring about significant improvements in farming practices and even entire communities (Duveskog et al., 2011). These programs are a great example of development practices that focus on empowerment and highlight the positive social sustainability INGOs can have. The instances of ALPS and FFS can reveal learning points for the INGOs in this study, especially the three I have identified as the most detailed about SS.

The Donor-INGO-Beneficiary System

Outlined in this paper's literature review was the process of systems thinking (Arnold & Wade, 2015). This form of thinking will be used to conclude this paper. In systems thinking, Arnold & Wade (2015) argue that understanding the framework of systems, improves "the capability of identifying and understanding systems, predicting their behaviors and devising modifications to them in order to produce desired effects" (Arnold & Wade, 2015, p.676).

A system is composed of three parts: (1) the elements, or the characteristics of the system, (2) the interconnections which are the ways the elements relate to each other, and (3) the purpose which is the reason why the system exists (Arnold & Wade, 2015). In the context of this paper, the donor-INGO-beneficiary relationship is a system in and of itself. In this system the donors, INGOs, and beneficiaries are its elements. These elements are highly interconnected in that they all play a role in the INGO work. Donors fund the INGO programs and services, INGOs manage the funds while trying to have the most effective impact on beneficiaries, and beneficiaries receive the INGO products or services. The absence of any of these elements means that the system is dysfunctional. So, with any of these elements missing from the system, there would be no INGO in the first place. Where does social sustainability fit in the donor-INGO-beneficiary system? Social sustainability, I argue, is the mechanism in which the donor-INGO-beneficiary system is regulated and controlled.

The purpose of this donor-INGO-beneficiary system can be widely interpreted. However, the general goal of this system is to reduce poverty and inequality among beneficiaries. Yet, the motive behind this goal is often obscured and unclear (Bennett, 2010; Bhati, 2021; Krause, 2014).

Though each element in this system is interconnected, some elements have a larger influence in the system than others. For example, donors usually have a larger say in the operations of the INGOs and in many

instances this influence has a negative impact on INGO work (Bennett, 2010; Brinkerhoff & Brinkerhoff, 2004; Jordan & van Tuijl, 2006; Krause, 2014). This is reflected in my finding that the vast majority of the INGOs in this study have websites catered to donors.

As for the INGO-beneficiary relational power dynamics, beneficiaries usually do not have as large of a say in the INGO's operations, but the findings in this paper reveal that INGOs seem to be trying to change this. This is seen through the various SS documents and the empirical data published by the INGOs to show how they are incorporating more beneficiary decision making into their programs. Likewise, the fact that all eight of this study's INGOs have separate publications on SS highlights that INGOs are actively discussing SS. SS text is easily available with a few clicks of the mouse on the majority of these INGOs' websites.

Just because INGOs are actively discussing and publishing SS material does not mean that they are living up to their claims. Nonetheless, it is clear that this topic is highly important for INGOs. Even if the primary purpose of publishing SS text was purely for marketing purposes, or to gain the trust of donors, the fact that some of the largest INGOs in the world seem to be modifying their accountability to shift more power to beneficiaries in the system is important to realize.

When trying to understand the donor-INGO-beneficiary system one must also acknowledge the fact that INGO conceptualizations of SS widely vary. As revealed in these findings, there seems to exist a spectrum of specificity regarding INGO text on SS. I interpret this spectrum as the degree in which INGOs place importance on the concept of SS as a whole. Similar to what I said in my analysis, if an INGO publishes thorough SS text and makes it easily available on its website, then it can be implied that the INGO places importance on the concept of SS and vice versa. So even though I argue that it appears that INGOs see significant importance on the concept of SS, the degree of importance varies.

Trends in INGO Conceptualizations of SS

As explained in the literature review, the concepts of sustainability, (including social sustainability), sustainable development, and even the INGO itself are all relatively new to society (Detraz, 2017; Gibson, 2019). This is reflected in my findings as well. Many of the SS texts studied in this paper are relatively new (often less than a decade old) and I found few examples of SS text going back more than a decade. Table 4 outlines some trends in SS text that I noticed in my research. Finding trends in SS was difficult. As a result, not every INGOs in this study are in this table. The only INGOs

included in table 4 are the ones that I was able to find relevant information for. My findings shown in Table 4 confirm the notion that SS is a newer concept. It appears that INGOs have only recently been publishing SS text. Because of the recency of this topic, conceptualizations of SS are and have been subject to consistent change, modification, adaptation. This paper is a snapshot of how certain INGOs conceptualize SS in a certain time frame.

Changes over time regarding SS discussions

Save The Children	Accountability reports began annual publications in 2014 (Save the Children International, 2021).
World Vision	Appears to have an accountability report dating back to 2009.
CRS	The earliest publication under the "Monitoring Evaluation Accountability and Learning" topic is from 2011 *(Research and Publications,* n.d.).
Plan International	Accountability reports were first published in 2007.

Table 4.

Theme 3 (Trends in Conceptualizations of SS) Notes

Conclusion

This paper has examined social sustainability in an INGO context while also examining the relational power dynamics between donors, INGOs, and beneficiaries. The goal of this paper is to understand how some of the largest Northern INGOS in the world conceptualize and discuss SS on their respective websites. I have done this by performing a qualitative content analysis of eight INGO websites. I have argued that my research is important because these INGOs are dealing with hundreds of millions of dollars in funds and consequently have significant influence on the communities they serve. Moreover, INGOs are significant actors in some of largest problems the world faces today. Because of these reasons, it is important to see where these INGOs are coming from and learning how they are managing their impacts on the communities they work in. What INGOs are saying about SS and how they understand SS can help academics, INGOs, and ordinary donors alike create an awareness of the social sustainability challenges INGOs face. Now that I have developed some understanding of how INGOs conceptualize SS and their relational power dynamics in my analysis section, I will conclude my findings by situating

them within the context of current literature and providing suggestions for future research.

With all of these findings in mind I will now succinctly answer my research question: The most common way that INGOs discuss SS in this study is through accountability reports which are inward looking documents. However, forward looking documents such as vision statements and strategy outlines are also common. The most common term used to conceptualize SS is the term accountability. The specificity of the INGOs' SS practices varies greatly between INGO though in the majority of cases SS text is not only accessible from the front page, but also all eight INGOs have separate publications on SS. The US sites do not appear to have much SS text for a variety of possible reasons. In terms of the power dynamics conveyed on the sites, it is clear that they are designed for donors and the mentioning of beneficiaries on the front page could be interpreted in a variety of ways. Finally, empirical data supporting the INGOs' SS claims varies in content and is infrequent. However, there are a number of INGOs in this study who have provided large amounts of texts on SS which include empirical data to support their SS practices.

With this understanding of the donor-INGO-beneficiary system in mind, researchers can, according to Arnold and Wade (2015), know how to modify and devise modifications "in order to produce desired effects" (Arnold & Wade, 2015, p.676). What are the desired effects of the donor-INGO-beneficiary system? I conclude that they are (1) an increased INGO awareness of the challenge of ensuring SS practices and (2) a strong intention for INGOs to publicly share information on their SS practices via their websites. Empirically showing how certain programs or systems are socially sustainable can be beneficial for all elements in the donor-INGO-beneficiary system. INGOs are moving in this direction, but there still seems to be a noticeable gap between SS programs and data supporting those programs.

Accountability reports, or inward-looking documents, seem to be the richest in SS text, and thus the most effective at communicating SS. Yet not even the majority of INGOs in this study have easily accessible accountability reports. With that being said, another idea INGOs could consider is to place increased importance on the publication of inward-looking accountability documents. Once again, this allows for transparency and challenges INGOs to examine their own social sustainability.

Content analysis of INGO media regarding SS is rare in existing literature. Many content analyses are focused on the presentation of INGO beneficiaries in INGO media as in the example of Bhati (2021). However,

there is not much scholarship that examines INGO media in an SS context. Future research could look more into this area. For example, I have mentioned broad trends in INGO website text regarding SS. However, this topic requires much more, in-depth research. As I have previously mentioned, this study is essentially a time stamp of how certain INGOs conceptualize SS. INGO websites are constantly changing and thus future research could look into these changes. Future research could also look into how donors themselves interact with INGO websites and what their views are on SS. Lastly, future research could look into SS text on southern INGO media and possibly even do a comparison between northern and southern INGOs. Each of these recommendations for future research are gaps in my own research. My research in this paper is aimed to provide a broad outlook on social sustainability and INGOs.

References

About ActionAid USA. (n.d.). ActionAidUSA.org. Retrieved March 29, 2024. https://www.actionaidusa.org/about/ .

Accountability and transparency. (n.d.). Policy-Practice.Oxfam.org. Retrieved on April 1, 2024. https://policy-practice.oxfam.org/accountability-transparency/

Ahmed, B. (2022, November 22). *Mali govt bans aid groups receiving funds from France.* AP News. https://apnews.com/article/europe-africa-france-mali-west-5 703cdf0a429f65143abf437a2028d84

Arnold, R. D. Wade, J. P. (2015). A definition of systems thinking: A systems approach. *Procedia Computer Science, 44, 669-678. https://doi.org/10.1016/j.procs.2015.03.050.*

Audet, F. (2015). From disaster relief to development assistance: Why simple solutions don't work. *International Journal, 70*(1), 110-118. https://doi.org/10.1177/0020702014562595

Bennett, S. (2010). *Structural incentives and the performance of international relief and development NGOs.* Washington University in St. Louis. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3dfcf5a3d3f7776d777 96a1480fc25b138df40bf

Bhati, A. (2023). Is the representation of beneficiaries by international nongovernmental organizations (INGOs) still pornographic? *Journal of Philanthropy and Marketing, 28*(4), e1722. https://doi.org/10.1002/nvsm.1722

Brinkerhoff, D. W. Brinkerhoff, J. M. (2004). Partnerships between international donors and non-governmental development organizations: Opportunities and constraints. *International Review of Administrative Sciences, 70*(2), *195-432.* https://doi.org/10.1177/0020852304044254

CARE International. (n.d.). CARE-International.org. Retrieved March 29, 2024 from https://www.care-international.org/

CARE International. (n.d.) *CARE's SDG impact.* CARE-International.org. https://www.care-international.org/sites/default/files/files/SDG_Impact_Report_190906.pdf.

CARE International. (2021). *Care 2030 vision: Harnessing collective power to fight poverty, and achieve social justice.* Care-International.org. https://www.care-international.org/sites/default/files/files/Vision_2030.pdf.

CARE International Secretariat. (2021). *CARE international impact, accountability & learning report 2021.* Care-International.org. https://www.care-international.org/sites/default/files/2022-12/CARE%20International_Im pact%2C%20Accountability%20%26%20Learning%20Report%202021.pdf

Catholic Relief Services. (2020). *CRS commitments to the 2030 Sustainable Development Goals. https://www.crs.org/sites/default/files/usops-resources/sdg_crs_goal_alignment.pdf*

Detraz, N. (2017). *Gender and the environment.* Polity Press.

D'Hollander, D., Marx, A., Wouters, J. (2013). *Integrating human rights in development policy: Mapping donor strategies and practices. https://dx.doi.org/10.2139/ssrn.2286204*

Duveskog, D., Friis-Hansen, E., Taylor, E. (2011). Farmer Field Schools in rural Kenya: A transformative learning experience. *Journal of Development Studies, 47*(10), 1529-1544. DOI: 10.1080/00220388.2011.561328

Esslemont, T. (2015, July 15). *Exclusive: Aid charities reluctant to reveal full scale of fraud*. Reuters. https://www.reuters.com/article/us-aid-business-fraud/exclusive-aid-charities-reluctant-to-reveal-full-scale-of-fraud-idUSKCN0PP00V20150715/

Eyben, R. (2006). *Relationships for aid*. Routledge.

Fowler, A. (1997). *Striking a balance*: *A guide to enhancing the effectiveness of non-governmental organisations in international development.* Earthscan.

Gibson, T. (2019). *Making aid agencies Work: Reconnecting INGOs with the people they serve.* Emerald Publishing Limited.

Gilbert, K. (2020). *Why well-meaning NGOs sometimes do more harm than good.* Insight.Kellog.Northwestern.edu. https://insight.kellogg.northwestern.edu/article/international-aid-development-ngos-crow ding-out-government

Hembling, J., Perrin, P., & Castleman, T. (2024). *Measuring respect for human dignity: Guidance for development and humanitarian programs*. https://www.crs.org/our-work-overseas/research-publications/measuring-respect-human-d ignity-among-project-participants

Hughes, P. M. (2004). *Case study of three international relief agencies: Ethics, humanitarian aid and complex political emergencies* (Order No. 3165278). Available from ProQuest Dissertations & Theses Global. (305212212). https://www.proquest.com/dissertations-theses/case-study-three-international-relief-agenc ies/docview/305212212/se-2

Human Rights Careers (n.d.). *About us.* HumanRightsCareers.com https://www.humanrightscareers.com/about/

Human Rights Careers (n.d.). *The fifteen biggest NGOs in the world.* https://www.humanrightscareers.com/issues/biggest-ngos-in-the-world/ HumanRightsCareers.com

International Rescue Committee. (2022). *2022 Annual report.* Rescue.org. https://www.rescue.org/sites/default/files/2023-07/AnnualReport_2022.pdf

Jordan, L. van Tuijl, P. (2006). *NGO accountability: Politics, principles and innovations.* Routledge.

Lewis, D., Kanji, N., Themudo, N. S. (2021). *Non-governmental organizations and development.* Routledge.

Lindenberg, M., Dobel, P. J. (1999). The challenges of globalization for Northern international relief and development NGOs. *Nonprofit and Voluntary Sector Quarterly, 28*(1), 4-24. *https://doi.org/10.1177/089976499773746401*

McEwan, C., Mawdsley, E. (2012). Trilateral Development Cooperation: power and politics in emerging aid relationships. *Development and Change, 43*(6), 1185-1209. *https://doi.org/10.1111/j.1467-7660.2012.01805.x*

Medecins Sans Frontieres. (n.d.). *What we do.* DoctorsWithoutBorders.org. https://www.doctorswithoutborders.org/what-we-do

Monika, K. (2014). *The good project: Humanitarian relief NGOs and the fragmentation of reason*. University of Chicago Press.

Munasinghe, M. (2009). *Sustainable development in practice: Sustainomics methodology and applications*. Cambridge University Press.

O'Leary, S. (2017). Grassroots accountability promises in Rights-Based Approaches to Development: The role of transformative monitoring and evaluation in NGOs. *Accounting, Organizations and Society, 63*, 21-41. *https://doi.org/10.1016/j.aos.2016.06.002.*

Our approach to sustainable change-how do we sustain improvements in child well-being, (n.d.). WVI.org. https://www.wvi.org/our-approaches/sustainable-change.

Oxfam. (2022). *Integrity at Oxfam.* oi-files-d8-prod.s3.eu-west-2.amazonaws.com. https://oi-files-d8-prod.s3.eu-west-2.amazonaws.com/s3fs-public/2023-01/Integrity%20R eport%202022.pdf

Plan International. (n.d.). Plan-International.org. Retrieved March 29, 2024 from https://plan-international.org/ .

Research and Publications. (n.d.). CRS.org. Retrieved April 26, 2024 from https://www.crs.org/our-work-overseas/research-publications.

Sánchez-Páramo, C., Hill, R., Mahler, D. G., Narayan, A., Yonzan, N. (2021, October 7). *COVID-19 Leaves a legacy of rising poverty and widening inequality.* Blogs.WorldBank.org. https://blogs.worldbank.org/en/developmenttalk/covid-19-leaves-legacy-rising-poverty-and-widening-inequality#:~:text=About%2097%20million%20more%20people,estimated%20to%20have%20been%20lost.

Save the Children. (n.d.). SavetheChildren.net. Retrieved March 29, 2024 from https://www.savethechildren.net/ .

Save the Children International. (2023). *Save the Children accountability report 2022: Strengthening our accountability in 2022.* ResourceCentre.SaveTheChildren.net. https://resourcecentre.savethechildren.net/pdf/Save-the-Children-Global-Accountability-Report-2022-Strengthening-our-accountability-in-2022.pdf/ .

Save the Children International. (2021). *Global accountability report 2020: Our journey of accountability.* ResourceCentre.SaveTheChildren.net. https://resourcecentre.savethechildren.net/pdf/STC-Global-Accountability-Report-2020-F INAL.pdf/

Scott-Villiers, P. (2002). The struggle for organisational change: How the ActionAid Accountability, Learning and Planning System emerged. *Development in Practice, 12*(3), 424-435. https://www.jstor.org/stable/pdf/4029512.pdf

Unerman, J., O'Dwyer, B. (2010). NGO accountability and sustainability issues in the changing global environment. *Public Management Review, 12*(4), 475-486. DOI: 10.1080/14719037.2010.496258.

United Nations Global Compact. (n.d.). *Social sustainability.* UNGlobalCompact.org. https://unglobalcompasystect.org/what-is-gc/our-work/social

White, M.D., & Marsh, E.E. (2006). Content analysis: A flexible methodology. *Library Trends 55*(1), 22-45. https://doi.org/10.1353/lib.2006.0053.

World Vision. (n.d.). WorldVision.org. Retrieved March 29, 2024 from https://www.worldvision.org/ .

World Vision International. (n.d.). WVI.org. Retrieved March 29, 2024 from https://www.wvi.org/ .

World Vision International. (March, 2023). *Global accountability report 2022.* WVI.org. https://www.wvi.org/sites/default/files/2023-04/WVI%20Accountability%20Report%202 022.pdf .

Karma Rakow graduated *summa cum laude* from the University of Memphis in December 2024 with a major in Anthropology and a double minor in English and International studies. She also earned Honors in Anthropology, University Honors with Thesis, and Undergraduate Research Scholar designations. Throughout her time at the University of Memphis, she has worked extensively under Dr. Lindsey Feldman and has presented at the Works in Progress Symposium. She has a passion for humanitarian projects and social justice and hopes to use her anthropology background to fight for human rights. Karma is currently pursuing an MA in Applied Anthropology from the University of Memphis, where she is researching immigration and refugee rights.

Karma Rakow

Drag Performers in the American North and South:
Identity and Resistance

Faculty Sponsor
Dr. Lindsey Feldman

Abstract

In 2023, the state of Tennessee passed the adult cabaret act—more colloquially known as the 'drag ban.' This law prevented so-called 'adult' performances from taking place in public, in the presence of children, within 1000 feet of a school, public park, or place of worship. Drag performances have always been political. Drag symbolizes queerness, serving as not only an open declaration, but a celebration of said queerness as an immutable aspect of self and foundation of personal identity. By subverting traditional notions of masculinity and femininity, drag performers destabilize the assumed naturalness of gender identity and expression. Utilizing Grand Rapids, Michigan, and Memphis, Tennessee as two operating field sites, this research takes a comparative ethnographic approach to analyze how the changing political landscapes in the American North and South affects the identities of drag performers.

Introduction

I sit next to a pair of photographers at a small round table in a black painted bar in Michigan. The dance floor in front of us has been cleared out and red curtains drape over the wall to mimic the effect of a theater show. To my right, YouTube clips from previous seasons of RuPaul's *Drag Race* play on a projector with the volume turned off, leaving us to guess what music the drag queens are dancing to. On my left, the photographers are setting up lighting and discussing the best way to work without getting in the way of the performance. Glasses clink behind me, and I look around to see the bartender laughing as he mixes a drink in a rainbow shaker. The audience seating is mostly empty—allowing me to sit in the front row and count cash to make sure I grabbed enough from the ATM for the night. A comfortable chatter fills the room as people slowly filter in, the conversations drowning out the music playing from the speakers behind the stage. Drag shows are a uniquely sensory experience. From the construction of the stage to the layout of the audience, even as an audience member, one's body is pulled into the show before it can even begin. Without warning, the lights go dim and the chatter comes to a halt as an oversized disco ball lowers from the ceiling. The first few notes of a song begin to play, and I hold back a laugh as a black drag queen in a shimmery silver cocktail dress takes the stage and begins to lip-sync Cher's cover of "Walking in Memphis." The irony hits me like a stiletto hits the floor. I am more than seven hundred miles away from Tennessee- a fourteen-hour drive away from the city that inspired this research- but in that moment I was able to see exactly how clearly Memphis and Grand Rapids were connected, and how it was drag performers that facilitated that connection.

Without drag queens there would be no queer rights movement (Rothenberg, 2021). In 1969 a chorus of drag queens combated a police raid of Greenwich's Stonewall Inn, making history as the literal and metaphorical first bricks were thrown in protest. In 2023, a straight man (who dressed in drag in college) passed a ban across the state of Tennessee that required all "adult cabaret" be removed from public performance venues where children might be present. A year later, one of the world's largest and most popular performers is a female drag queen named Chappell Roan who brings local drag performers into her shows. Drag performers and performances hold a complicated place in current cultural discourse, as demonstrated by the above examples. Both lauded and criticized by national audiences, drag performers and performances are inherent representatives of queerness (Newton, 1979), symbolizing queer history and

subversion of traditional gender norms. It was this state of contemporary cultural discourse that inspired the course of this research.

For the purposes of this thesis, I define drag performers as individuals of any gender who adopt an exaggerated gendered persona for the purpose of entertaining audiences. These personas are physically displayed through costume choice, makeup, and styles of dance. While the most common impression of drag performers is that of cisgender gay men dressed in an overly feminine style (wigs, heels, and showgirl dresses), this research studied all performers and performances, regardless of the gender (perceived or otherwise) of the performer. With exceptions (such as during pride month and for specially booked events), drag is most often performed in queer bars. Queer bars contain a special meaning for those in the LGBTQ+ community, often serving as safe havens for those who cannot openly express non-normative sexual or gendered identities. Given the cultural meaning of the queer bar, they serve as key sites for the development of collective identity among drag performers.

There is a plethora of existing research surrounding drag performers and identity (see, Butler 1993, Taylor et. Al, 2015, Kaminsky & Taylor, 2008), however, there is very little existing research that discusses drag performers in a national context. In this thesis, I argue that the 2023 Tennessee drag ban (which will be discussed later) served as an inciting event for the creation of a collective drag identity among performers in the American North and South. I further argue that discourse and political rhetoric intersected and affected the ways in which performers in both geopolitical environments perceived themselves as individuals and as members of a drag community. My original research questions were: How do actively changing laws affect drag queens on personal, social, and economic levels? How do these laws affect the way they perceive themselves? Do performers see drag primarily as an identity or a source of income? Does drag performance shape the way that one views their own gender? Resistance was not among my original questions, but it emerged as a strong theme throughout my research that I believe to be essential in discussing the identities of performers. While not every performance is explicitly political, when one's identity as a performer is being threatened, performance becomes a political act that asserts an individual's personal identity while situating them as belonging to a collective.

Divided into four sections, I discuss identity and its connection to resistance among drag performers by comparing two main research sites: Grand Rapids, Michigan; and Memphis, Tennessee. The first section introduces previous theoretical literature surrounding drag performers, gender,

collective identity, and performance. The second section discusses the anthropological methods used throughout the fieldwork and research process. The third section discusses the ways in which drag performers created a collective drag identity in reaction to changing laws that sought to prohibit drag performances, while at the same time asserting their own personal racial and gendered identities to operate as resistance on an individual level. The thesis concludes with a discussion of the collective identity established by drag performers, and implications for the future of drag.

Literature Review

Drag and its performers do not exist in a historical vacuum. The first known drag performer in American history was a former slave by the name of William Dorsey Swann (Shane, 2023). In the late 1880s in Washington DC, Swann created what are now considered to be the first drag balls. These balls– held in secret due to the fear of police raids– included cross-dressing, dancing, and competitions. In mimicry of the antebellum era "cakewalks," where black dancers would mimic the attitudes and mannerisms of white men and compete to win cake; Swann's resistance dances allowed performers the chance to be crowned "queen" for the evening. Swann died in 1925, but his drag balls have lived on to the modern era, visible in vogue halls and drag shows every week at local bars across the country. From its conception, drag has been an act of protest by queer people and people of color, and it has been a key site in the formation and development of individual identity, collective identity, and queer culture through performance. As the central questions surrounding my research were located around the theme of identity, I utilized a historical literature review in combination with feminist theory, queer theory, and performance theory to argue how sites of political contention serve as a foundational site for the formation of a collective identity among drag performers.

Collective identity, according to Cristina Fominaya, is centered around collective action. Collective identity is a process that involves "cognitive definitions about ends, means, and the field of action" (Fominaya, 2010). This is central to my examination of collective identity, as it allows for variation between the motivations of individual actors. Individuals do not necessarily have to agree on ideologies, beliefs, interests, or goals to come together in the name of collective action for a social movement. Furthermore, when collective identity is created in social movements, it tends to define itself in opposition to dominant cultural practices. A key aspect of collective identity- especially as used in relation to drag performers- is that there is significant overlap between personal and social identities and

collective identity. When it is identity itself that is being threatened by political and legal forces, this becomes essential to understanding the way that identity is formed among drag performers.

Identity is inherently political- queer identity even more so. Esther Newton first brought drag into the anthropological arena with her book "Mother Camp: Female Impersonators in America." Written a mere ten years after the Stonewall riots, Newton spent two years conducting field-work on drag queens as they performed and lived their daily lives. Newton's female impersonator is the glamorous and melancholy urbanite-disillusioned, cynical, and yet still on stage every other night. For these queens "exploitative sexuality is the root and framework of all human motivation" (Newton, 1979) and the polar division between male and female is the most fundamental distinction in human existence. Newton locates the formation of collective identity among her female impersonators in oppression and collective longing. These female impersonators are, inherently, impersonators. They perceive themselves as morally and sexually polluted and– understanding themselves to be lacking in comparison to the masculine ideal– seek fulfillment through performative Hollywood-style high-femininity. While "Mother Camp" may have been revolutionary for its time, in the 45 years since it was written, drag has left the underground dive bars and joined the national conversation. While I disagree with Newton's conclusion of drag performers as melancholic of their queerness, I do agree with her finding that collective identity among drag performers emerges from situations of oppression.

Drawing on their ethnography conducted at the 801 bar in the Florida Keys (Rupp & Taylor, 2015) Rupp, Taylor and Gamson (2015) explore collective identity among Florida drag queens as formed and displayed through acts of protests. Their findings "illuminate the way cultural performances can be used as political tactics: their role in contesting the dominant order, the degree of intentionality involved, and the kinds of collective identity work they embody." They expand on this to claim that collective identity is created among drag queens through deliberately political performances. Drawing on Newton, the authors theorize that the body of the drag queens as performance artists is used to draw attention to the social construction of gender and sexuality while contesting dominant gender roles. In this way, collective identity among drag performers is embodied through the act of performance. As identity and performance are both political, the body becomes a site of both protest and politics.

It is nearly impossible to discuss identity and performance without mentioning Judith Butler's theory of gender performativity. Butler's theory is a foundational theoretical grounding in feminist and queer studies, providing a philosophical and theoretical basis for understanding gender as it is expressed, embodied, and enacted in daily life. Gender performativity- according to Butler- is the way that one enacts gender through their choices, movements, bodily decor, etc, through a "stylized repetition of acts" (Butler, 2017). Gender is thus constituted of a series of conscious and unconscious efforts conducted by an individual in accordance with (or in direct opposition to) the social norms of appropriate gender expression.

> "My suggestion is that the body becomes its gender through a series of acts which are renewed, revised, and consolidated through time. From a feminist point of view, one might try to reconceive the gendered body as the legacy of sedimented acts rather than a predetermined or foreclosed structure, essence, or fact, whether natural or linguistic" (Butler, 2017).

Butler cites their theory as inherently political, as gender is a site upon which cultural power is exerted and norms are enforced. In their later piece "Gender Trouble," Butler elaborates on manners in which performativity and drag intertwine to show the lack of connection between gender and sex. Responding directly to Newton's "Mother Camp," Butler claims that drag as a literal performance capitalizes upon the dissonance of bodily autonomy of the performer and the gender being performed. Drag thus demonstrates the imitative structure of gender and the acts that constitute it (Butler, 1999). "If the anatomy of the performer is already distinct from the gender of the performer," Butler says, "and both of those are distinct from the gender of the performance, then the performance suggests a dissonance not only between sex and performance, but sex and gender and sex and performance" (Butler, 1999). This theory was essential to the development of my research questions and my approach to understanding identity. I utilize Butler's theories to discuss both individual and collective identity among performers, as understanding performativity as different from– but related to– literal performance. This theory allows one to understand the importance of gender performativity for those of non-normative gender identities in the development of one's personal identity, but also allows for an understanding of how performativity connects with performance to help create collective identity.

Marlon Bailey delves into queer performance theory in his 2013 ethnography, "Butch Queens Up in Pumps: Heder, Performance, and Ball-

room Culture in Detroit." While not studying drag performers, the location of the ethnography (a mere three hours away from where my own fieldwork was conducted), the theoretical basis of performance and gender theory, and the significance of cultural history embodied through dance (ie. conscious performative acts) were essential in the creation of my fieldwork and the construction of my findings. Bailey's work dives deep into the world of ballroom culture first made visible to the world in Jennie Livingston's 1990 documentary, "Paris is Burning." Ballroom culture is primarily made up of queer BIPOC, and as such, ethnic, cultural, gendered, and sexual identity are present in every aspect of an individual's performance. As Bailey puts it, "Performance makes it possible to revise, negotiate, and reconstitute gender and sexual roles and relations while constructing a more open minoritarian social sphere" (Bailey, 2013).

Vogue dancing– consisting of fluid, whole-body movements including arm control, body awareness, and understanding of musical beats, rhythm, and syncopation– is considered to be one of the most important aspects of ballroom culture. Emerging from African diasporic dance traditions in combination with house and hip-hop music, vogue dance symbolizes racial and cultural identity through its music and movement. However, it also produces new forms of collective identity through its continued use in ballroom culture and its spread into the wider world as a recognized form of dance. Having been appropriated from ballroom, vogue is often used in drag performances as a "showstopper," containing some of the most dynamic dance moves and the best potential for audience interaction. While not often recognized by non-POC audience members, vogue as used in drag often displays the cultural and historical connections between ballroom, BIPOC, and drag. As such, when utilized in drag, vogue dance serves as a way in which collective history and identity is embodied and reproduced through conscious performative acts.

Using the above authors in conversation, I explore the ways that individual identity is constitutive of collective identity by analyzing the ways that drag performers assert their own non-normative gendered, racial, and cultural identities through literal performances on stage to serve as a quotidian act of resistance that assists in the formation of a national drag identity in the face of political and legal contestation.

Methods

The original aim of my thesis and the construction of this research was to understand how changing laws that directly targeted drag performers affected the formation of identity of said performers by comparing two

sites in the American North and South. This research was conducted over the course of three months, from June to August 2024, in Grand Rapids, Michigan, and Memphis, Tennessee. The respective cities were chosen for both their political uniqueness and my own personal connection to each location. Grand Rapids and Memphis present interesting research sites due to their political contrast to the states they are situated in.

Grand Rapids is a relatively conservative city in one of the "blue wall" states, where Memphis is often called the "blue dot" of conservative Tennessee. It was members of the Memphis-based drag theater troupe "Friends of George" who were responsible for suing the state of Tennessee for the drag ban that occurred in 2023 (Kashiwagi, 2023), making the location more than suitable for fieldwork surrounding drag and identity politics. Grand Rapids is situated in a political context where the governor of the state has followed in the opposite direction of Tennessee's current governor, and has signed into state policy a series of laws that protect LGBTQ+ individuals from discrimination (HRC Staff, 2023). As such, Grand Rapids made for a fascinating environment from which to conduct comparative analysis. Given the geopolitics of each city, they served as ideal sites for research surrounding politics and collective identity.

Following in the tradition of anthropology, my fieldwork took the form of an introductory-style participant observation. Participant observation is a method in which the researcher actively participates in the culture or community they are studying, thus allowing them to gain both an inside and outside perspective on the way individual cultures function (Bernard, 2017). Participant observation allows for a more subjective approach to qualitative data collection and analysis, as the positionality of the researcher is something explored in both the research as well as the final product. While a true participatory research model would have required me to perform in drag myself, the time limit in place for this project as well as my own unfamiliarity with this model of fieldwork limited me from that aspect of participant observation. As such, my participation and engagement with performers was mostly from the perspective of an audience member. This allowed me to examine interactions between performers and audience members while shows were occurring, and to build rapport with the performers by becoming a regular attendee.

I was not sure what to expect from my first night in the field. Despite having grown up in Grand Rapids, I had only ever heard of this bar from older queer people. Rumors Nightclub is Grand Rapids' oldest queer bar and one of the only places where drag shows occur on a weekly basis. Given the historical importance of Rumors to the local LGBT+ community,

it served as my primary research site while in Michigan. I would arrive approximately twenty to thirty minutes before the beginning of the show so I could get seating close to the stage and speak with bar regulars while waiting for the show to begin. Despite being the oldest queer nightclub in the city, Rumors is not large. Consisting of a bar (decorated in rainbow bottles of Absolut Vodka), 15 small tables, a projector beaming RuPaul reruns and Lady Gaga concerts, and a mirrored wall decorated in red curtains before every show; the space is constructed so that people only have two things to look at upon entering: the stage and the bar. The construction of the space is designed to promote audience interaction with performers and performances, as the stage is visible from every angle of the club. Shows at Rumors run from 10pm to 12am, with an open dance floor from midnight to 3am where a DJ in drag plays pop music until closing time.

When in Memphis, I conducted the majority of my fieldwork at Atomic Rose—the only queer bar on Beale Street. The show start times would vary each week, but usually starting between 10pm and 11pm, but always ending by midnight. Drag performers could be found at the bar before shows in full getup, and were more than willing to take pictures with tourists and bachelorette parties stopping by Beale for the night. As far as I could tell, Atomic Rose had few regulars, and instead relied on the Beale Street bar-hopping tradition to draw in audience members each week. The nightclub is rather large, combining a bar-style restaurant with a chic nightclub. There are two bars where patrons can order drinks, multiple sets of tables, and a T-shaped stage with two steps that lead down to the permanent dance floor. Paintings of queer icons such as RuPaul, Ariana Grande, Selena Gomez, John Lennon, and Gloria Gaynor line the dark-painted walls. However, the real attention-grabber is the large golden throne at the entrance. With a multicolored spotlight always aimed at the chair, it is common to see both performers and guests resting on the throne to watch the show.

I conducted four semi-formal interviews over the course of three months. As drag is a politically sensitive topic (especially since my research took place in the months prior to the 2024 US election) my interlocutors were given the option whether to allow their personal information to be used in the course of the research prior to the start of their interviews. They were encouraged to ask questions off the record prior to and immediately following the interview, and they were also given copies of the recording transcript to edit or comment on as they felt necessary. Importantly, the two performers from Tennessee requested I use pseudonyms rather than their real performance names. Due to the contentious legal status of

drag in Tennessee, it was important to both myself as a researcher and to my interlocutors that their safety was protected above all else. Both interviewees from Grand Rapids requested their real performance names be used in anything related to this research. The questions for the semi-formal interview were based off of a pre-written, IRB approved interview guide, but said questions were often altered based on the answers each participant gave. As such, the interviews often took the form of a conversation, with three out of the four interviews lasting for over an hour. Due to the fact that I was frequently travelling between states and all my interlocutors had one (or more) jobs outside of performing, three of the interviews were conducted over a password-locked zoom call, with the singular in-person interview taking place in a location of the interlocutor's preference.

It should be noted that personal pronouns are a complicated matter to cover when working with performers, as they often use different pronouns when on and off stage. To respect the privacy and the individual gender identification of my interlocutors and research participants, I will be referring to the individuals by the pronouns they prefer to use while in drag.

Lastly, the term "community" is utilized multiple times throughout this piece in relation to the queer community and drag community. Drawing on Neal Carnes, I define "community" as a people connected through conscious solidarity where deeply shared identity serves to bind people together (Carnes 2019).

Discussion

It is 9:30 at night and the show has yet to start. I made a mistake, believing that the show started earlier in the night than it did, and now I have no other choice but to wait until the audience filters into the club. It is the middle of Pride Month and, perhaps more importantly, only a year since the drag ban was repealed in Tennessee. There is no question that there will be an audience tonight. The only question is a matter of when they will get here. The air is filled with the scent of smoke, fries, and sweat, and pop music blares through the jumbo speakers on either side of the stage. I approach the bar and come to stand next to two towering drag queens, working up the nerve to introduce myself. Zahara Dessert wears a Vegas-showgirl style glittery dress and downs a shot as I speak to Brooklyn Oldshore, a queen with a large-than-life white wig and a polka dot dress. Zahara turns her focus to me, her makeup-accentuated gaze intense as I ask about what they thought about last year's drag ban and the new series of anti-transgender bills being proposed in Tennessee's courts. Brooklyn paused before answering,

"I hope," she said, looking away from me to wave at the bartender, "that in a few years we will be able to laugh about all this." On hearing Brooklyn's answer, Zahara's nose scrunched up, and a look of unease came over her. "What?" Brooklyn asked, turning to face Zahara fully, "you don't think so?"

Zahara's voice was quiet as she answered, "I don't think I can agree with you here."

Season 15 of RuPaul's *Drag Race* ended just as the 2023 Tennessee Senate Bill 662 and House Bill 30- known as the "Adult Cabaret Act" or "Drag Ban" was passed into state law. In the season finale, RuPaul Charles, host, judge, and founder of RuPaul's drag race announced a fund-raising partnership with the ACLU and the creation of a Drag Defense Fund- all dedicated to combating the state of Tennessee in defense of drag performers across the nation (Silver, 2023). While the ban was repealed mere months after it was signed into law- having been deemed unconstitutional by the courts- the effects of the ban linger in the drag community today. As demonstrated by the introductory vignette, individual performers vary on whether the ban set a precedent for future lawmaking practice both in the state of Tennessee and on a national scale. However, simply because individual actors disagree on the long-term effects of the ban does not mean that the drag community is divided. My first finding indicates that oppressive laws served as an inciting event for the creation of a collective drag identity through cultural discourse and political performances as acts of protest.

A year after the drag ban was brought down in court, Polara Polloi hosted a family friendly drag show in Memphis to celebrate Pride Month. Dressed in a rainbow tutu, Polara sat in the front of the room with her son on her lap as she read aloud a story about a child realizing they were transgender. As a burlesque dancer, drag queen, showrunner, and a mother, Polara was eager to participate in this project. Polara is the primary showrunner of a local Memphis drag troupe (whose name I have omitted for the sake of confidentiality) and thus is in charge of not only arranging shows with various venues, but also with organizing performers. During the 2023 drag ban, Polara decided to organize a family-friendly drag show to protest and challenge the idea present among Tennessee legislators that drag was inherently unsuitable for children. When asked about this show, Polara commented, "people [were] like, you shouldn't do drag in front of kids. And I'm like, well, I do. And it is perfectly safe, I promise."

Polara does not see the expression of her gendered identity as a performer as one that is inappropriate or unsuitable for young audiences, and

instead of retreating under political scrutiny, Polara decided to rally among her community and unite them through a protest event- despite knowing the possible consequences for hosting said event. In Polara's case, collective identity was created and reinforced through deliberately political performances. Collective identity is often centered upon direct engagement, and as such, action tactics such as protest shows create a site in which collective identity can be centered (Fominaya, 2010). This performance- to use Butler's terms- served to demonstrate the performative and political nature of gender by literally placing it on a stage and creating collective assertions of non-normative gender expression as an act of protest. This demonstrated collective identity by uniting performances through the ideology of drag and expressions of varying gender performances being suitable to be seen and understood by younger audiences- directly challenging the very nature of the Tennessee drag ban.

Seven hundred miles and two months away in Grand Rapids, I sit in a bar with a drag king named Cocky Divine. A black trans man from Detroit, Cocky is in love with both his drag and his community. He can be found most Sunday nights performing a slow strip-show at Rumors Nightclub, where he proudly reveals his top-surgery scars to an ecstatic crowd. He drinks wine and I snack on popcorn as he talks about the fear and anger he felt upon learning that members of his community were being threatened in Tennessee. "At one point," Cocky said, "we all wanted to go down to Tennessee and march in drag and be there […] Because these people [drag performers] aren't doing anything in their lives but living their life like you are. […] I was nervous for those people there, and what the bill meant for everybody else."

What this quote demonstrates is that Cocky- and other drag performers- feel a sense of community and connection with performers across the country. This sense of community was heightened with the threat posed against drag performers and encouraged the formation of a collective identity to serve as a site upon which protests could be built.

Kaminsky and Taylor (2008) claim that "Collective identity is negotiated over the life course of a movement in response to changes in the political context." Not only does it define the collective, it acknowledges some sort of injustice that has been committed against them, and attributes the blame to structural causes. What I found throughout the course of my fieldwork was that collective identity among drag performers was created by threats of political oppression—as demonstrated by Cocky's desire to mobilize drag performers—but collective identity was maintained through community-oriented political action.

Fundraising performances were a common occurrence in both Grand Rapids and Memphis. These shows demonstrated alignment with both the local and national LGBTQ+ community through their choice of performers and the organizations who received the donations after the shows. In Grand Rapids, a common fund-raiser was "Kamikaze night." This performance invited both amateur and professional drag performers to lip-sync, dance, and compete. These competitions mimicked RuPaul's *Drag Race*'s "lip-sync for your life," where performers must dance and lip-sync to win tips from the crowd. Whoever received the most tips at the end of the night would be deemed the winner of the competition. More importantly though, all the tips given to the performers throughout the course of the fund-raiser were donated to charity at the end of the night. These fund-raising shows serve to reveal to the audience members the current threats facing the LGBTQ+ community (and thus, the threats facing the performers themselves) and simultaneously operate as a form of identity expression and community acknowledgment. These types of performances further serve as a strategy to create feelings of belonging and collectivity among sympathizers and potential allies in the audience (Rupp & Taylor, 2008). Furthermore these performances are "staged to solicit strong emotion and are experienced by the audience as play rather than as serious, which allows them to attract participants who might not otherwise attend a political event" (Taylor, Rupp, & Gamson, 2015).

The time period in which this research was conducted is not irrelevant. My fieldwork took place in the months before the United States election, and as such, politics were often brought into the field whether in performances or in interviews. During the Grand Rapids Pride festival, a drag queen brought a democratic political candidate running for senate to the main stage and danced with him. The same candidate was brought up later that week in a show where the drag performers encouraged the audience to vote by creating a call and response of "Vote! Drag! Vote!" In both Michigan and Tennessee, the drag ban was used as an example of what would happen if Democrats did not win the election. In these discussions, queens would often use the word "we" when speaking. "I may not like Joe Biden," a queen commented at a performance in a public park, "but he's the best chance *we've* got. So go vote this November!" (emphasis added). In this case, the use of the word "we" indicated not just a conversation between the audience and the performing queen, but also an alignment and solidarity with drag performers across the country. These drag performances forge collective identity among performers "through the articulation of boundaries and community among those engaged in the performance," (Taylor, Rupp, and Gamson, 2015).

While collective identity was formed through alliance with specific political causes, it was also demonstrated through the active disavowal of political actors who were perceived to threaten the LGBTQ+ community and those who aligned themselves with said actors. When one audience member at Rumors expressed their conservative political affiliation, the evening's showrunner yelled into the audience, "Donald Trump is a damn fool. A racist fool. A misogynistic fool. A -phobic fool. A damn fool- and we all hate you." This statement was followed by cheering from the crowd as the offending audience member left the premises. Anger and fear were present in every interview I conducted with drag performers, however, the geopolitical context of each performer's primary location changed which emotion was most present. Performers in Grand Rapids (like the example above) demonstrated high levels of anger and indignity, where Memphis performers who were active in the drag scene during the 2023 drag ban expressed fear over the potential future of drag. As Oxy Daze, a non-binary performer based out of Memphis explained,

> "It was very scary because I had just found a site for myself. […] I had grown up very conservative and religious, and I had found a safe place for myself. And I thought it was gonna get taken away. […] It was scary to perform during that time, like I was going to have a target on my back just because I am who I am."

Collective identity defines itself in contrast to its relationship with other in the specific context in which they exist (Fominaya, 2010). When it comes to politically targeted minority groups such as drag performers, anger and fear provide a deeper basis for the formation and consolidation of group identity and solidarity. Emotions are essential in using collective identity as a basis for action, as emotional ties between actors can help them maintain their involvement in said actions, and help them overcome the effects of oppression (Fominaya, 2010). It is these emotions that serve as a functional starting point for the utilization of identity as a form of collective and individual resistance.

My second finding centers around the individual identities of drag performers and their connection to resistance. Where above I discuss the manners in which collective identity is formed through resistant narratives and actions, I now discuss how the assertion of drag performers' individual identities serve as quotidian acts of resistance. Drag operates as a microcosm of wider cultural norms, but in the face of political contestation, assertions and presentations of non-normative identities serve to establish said presentations as acts of protest. Drag is a political mode of critical performance where the categories of race and class intersect with gender and performativity through performance (Greene, 2021). These performances

solicit strong emotions among audience members in order to draw them into the world of drag-as-politics. At the same time, these performances are heightened by the individual performers, who assert their personal racial, cultural, and gendered identities during their performances as modes of self-validation and acts of resistance against those who would seek to regulate said non-normative identities.

Throughout my fieldwork, interlocutors expressed that their drag personas could not be separated from their individual, non-performing selves.

> "Building a drag persona is less contrived than it might appear. You don't make a mood board or cut and stick your idealized drag into your scrapbook. It's much more natural, a kind of slow unfurling of all the references you adored as a kid, a teen, an adult; an expansion and performance of all the things you are desperate to be," (Rasmussen & Rasmussen, 2020).

Every drag performer I spoke with made a similar claim to Rasmussen when asked about their drag personas. To them, their drag personas were not brand new, completely independent characters, but extensions and exaggerations of who they already are. However, these individual identities were displayed differently by performers of different racial, cultural, and gendered identities. These individual identities intersect in complicated and multifaceted ways, and serve as sites for both individual acts of resistance and for the development of collective identity among performers.

Ethnic and cultural identity are commonly displayed in drag performances through makeup, clothing, song choice, and dance styles. Mimi, known as the Xicana goddess of Grand Rapids, is a transgender Mexican woman. A fashion student, Mimi makes her own clothing, often copying flamenco and salsa-style dresses that allow her to incorporate traditional Latin American dances into her performances. Mimi's songs of choice often demonstrate her pride in her Mexican heritage, as she chooses classic Spanish love ballads (and once, notably, the Macarena). Mimi will often encourage audience members to sing in Spanish as she dances both on stage and among crowd members. In these performances, Mimi combines her gendered and cultural identities in what Butler would call a series of consciously performative acts that serve to demonstrate the complex nature of non-normative identities among drag performers.

Introduced on stage by Brooklyn Oldshore as the "epitome of drag," Geneva Convent is a Black drag queen in Memphis. Geneva is the type of glamor-drag goddess one would expect to see on RuPaul's *Drag Race*—or as a face in a beauty pageant. Geneva is always seen in a full-length and

bejeweled bodycon ballgown, earrings larger than the palm of a hand, and an opulent curly wig that goes down to her thighs. She does not dance, but instead saunters slowly around the dance floor, lip-syncing to Whitney Houston, Donna Summer, and R&B classics.

According to Marlon Bailey, drag and other gender-bending performances are fundamental to Black LGBTQ+ cultures. Bailey identifies the most important part of Black drag as recognizability. Performances must be recognizable by members of the Black community to be appreciated, and as such, impersonations of Black female singers and divas (such as Geneva's imitations of Whitney Houston) are enjoyed because they are understood and adored by Black audience members.

Both these examples demonstrate alignment with the performer's personal racial and cultural identities that are affirmed through their performances and interactions with their audiences (Bailey, 2013). Both Mimi and Geneva serve as examples for a common theme found throughout the course of my fieldwork. The drag performed by people of color serves various purposes. It challenges wider cultural norms of race (Tucket, 2023) and– in the case of my interlocutors– operates as a platform to manifest their voices politically. The drag personas of both Mimi and Geneva display similar levels of pride in their statuses as people of color.

However, in the midst of a culture where a black woman is running for president, performances that demonstrate pride in one's racial and ethnic heritage operate as an identity level of resistance to racist and sexist political and cultural discourse.

It would be a mistake to discuss the individual identities of drag performers without discussing the relationship between gender, performance, and performativity as they are related to protest. These frameworks influenced findings throughout my fieldwork and research by demonstrating not only the fluid nature of gender identification among performers, but also by creating the idea of protest and resistance as intimately connected to the daily lives of those with non-normative gender identities.

In 2023, US legislators introduced over four-hundred bills targeting LGBTQ+ (but mainly transgender) individuals (Butler, 2024). In these bills, the words "gender" and "gender ideology" can be found in most of them. 2023 saw Tennessee with two such bills: the aforementioned drag ban, and SB1– a bill that prohibits individuals under the age of eighteen from receiving gender-affirming care (Branstetter, 2024). SB1 is currently being reviewed by the Supreme Court, which will establish whether the limits placed by the state of Tennessee violate the constitution under the

Equal protections clause. In states where gender-affirming care is restricted, demonstrations of non-normative gender identities operate as public declarations of resistance against structures of oppressive power. In the national context, where access to gender-affirming care is a site of contention among legislators, these same displays represent open challenges to what Butler (2024) calls "abusive laws."

Cocky Divine makes his transness evident in every performance. Whether through the reveal of top-surgery scars, through outright declaration, or through symbolic acts– such as bringing a transgender flag on stage– there is never a question as to Cocky's identity. Cocky's drag is not always explicitly political, but his individual identity as a Black transgender man is so heavily politicized that any action he does can be interpreted as a performative act of protest. For Cocky, performance and gender performativity are so connected as to be nearly indistinguishable from one another. Cocky's drag thus constitutes a "complex and fluid act of identity expression where male and drag personas overlap and some aspects of performance are experienced as more natural than others," (Greene, 2021).

Oxy Daze made similar claims when asked about her drag persona:

> "When I'm performing, I don't like to have any sort of gender expression. I don't have to be like: this part of Oxy is male, this part is female. I'm a drag king. I'm a drag queen. I just say I do drag and I dress how [whatever] inspires me and I perform. So through drag I kind of found a way to perform and be queer at the same time."

While not all drag performers are transgender, all drag performances are displays of gender performativity. Butler claims that gender as expressed by drag performers suggests a dissonance between sex and gender and sex and performance" (Butler, 1999). For both Cocky and Oxy, it is their daily existence that is being challenged by the wave of anti-LGBTQ+ laws, and as such, demonstrating their gendered identities through conscious acts of performance serves as examples of both resistance and protest at the identity level.

Outside of drag, Cherry Poppins identifies as a cisgender gay man. "My big thing is," she said in an interview, "I want to look like a drag queen, not necessarily a real woman. My drag is female presenting, but I don't want people to think "that could be a man, could be a woman." I want to look like a drag queen." Cherry's drag serves to demonstrate the imitative and inherently performative nature of gender. By stating that she wants to be perceived not as a real woman, but as a drag queen, Cherry demonstrates another layer of the complicated intersections between

gender and performance. Cherry's drag reflects Newton's assertion that the drag queen symbolizes an open declaration and celebration of homosexuality, as well as Butler's central claim that drag reveals the inherently imitative nature of gender.

Despite the differences in their gender identities, all three of the individuals in the above examples identify as drag performers. In each case, the performer's drag personas were shaped by their intersecting racial, cultural, and gendered identities, but they identified as part of a collective that was threatened by the increase in oppressive laws that would target members of their community across the nation. Bailey (2013) theorizes performances as a form of cultural labor, in which performances operate as work that produces culture as a product that is then reaffirmed through community engagement. In attempting to regulate drag performers in Tennessee, legislators inadvertently assisted in creating a national collective drag identity that was centered in resistance through performances of non-normative identities. These performances served as cultural labor to create resistance and protest as key aspects of drag.

Conclusion

A notable limitation of this research was the sample size of the research participants. Due to fear of the changing political status of the state of Tennessee, many performers expressed hesitancy over participating in any sort of research project that draws attention to them as individuals. Safety was a large concern for my interlocutors, and in an election year, many were not certain that their words would not come back to haunt them later. As such, there was only an exceptionally brave few who were willing to meet with me for interviews. Furthermore, more time in the field would have been beneficial for this project, instead of constraining participant observation to three months between academic semesters. Many performers I spoke to were interested in participating in this research, but were limited in their capabilities due to the heightened demand for drag performances during the summer months when this research took place (Pride Month and July 4th are exceedingly busy times for drag performers).

Lastly—as mentioned in the introduction of this piece—there is very little anthropological work on drag performers and collective identity, and there is none that I could find that takes a comparative approach to collective identity of drag on a national scale. Future research on this topic should consider the way wider political challenges affect the way that drag performers perceive the state of their community. Further research should

also examine how identification and connection with a specific location affects the way that performers perceive their relationships between themselves, their communities, and the law.

Despite the repeal of the drag ban in 2023, the collective identity among drag performers was maintained across both field sites, allowing performers to identify each other as part of a community who- even if they held disagreements- were united against threats of censorship and political oppression that would see their labor and their community regulated. Drag performers do not exist in a cultural and historical vacuum. They are not a singular subculture existing outside of heteronormative culture. One cannot separate a performer's drag persona from their independent and multifaceted selves.

Identity is political. Queer identity even more so. The results of this study indicate that the identities of individual drag performers intersect on multiple levels to function as quotidian acts of resistance to oppression in the face of political discourse that would threaten their very existence. In the context of the 2023 Tennessee drag ban, the individual identities of performers helped make up a national collective identity of drag performers that is centered in resistance. This community culture of resistance consisted of acts of outright protest: such as organizing shows in direct opposition of the law, speaking at the Tennessee Capitol in defense of drag, and mobilizing national support through social media. However, for many drag performers, resistance also consisted of refusing to change their lifestyles and identities and instead choosing to continue existing in and out of drag as symbols of presence and of the queer community. While not every drag performance is explicitly political, when it is one's lifestyle and identity as a performer that is being threatened by forces outside of their control, performance becomes a political act that asserts an individual's personal identity while situating them among a collective.

This thesis demonstrates that collective identity among drag performers is both a process that is negotiated by individual performers and their personal identities, but also a product of political oppression that creates a collective drag identity. This thesis also demonstrates that identity serves as a site upon which resistance and protest is built. Despite their individual differences, drag performers across the country were able to use their identities as performers as common ground to demonstrate support for individuals in Tennessee who were being oppressed. Most importantly, this thesis demonstrates that the attempts by the state of Tennessee to ban drag performances did not work. Drag performers are still active and present

in Tennessee, and in the state's attempt to repress drag, they inadvertently created a strong sense of community and collective identification among performers across the country.

Works Cited

Bailey, M. M. *Butch Queens Up in Pumps: Gender, Performance, and Ballroom Culture in Detroit*. University of Michigan Press, 2013.

Bernard, H.R. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. 6th ed., Rowman & Littlefield Publishing, 2017.

Branstetter, G. "The Supreme Court case on Trans Health Care, explained: ACLU." *American Civil Liberties Union*. September 12, 2024, https://www.aclu.org/news/lgbtq-rights/the-supreme-court-case-on-trans-health-care-explained.

Butler, J. "Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory." *Readings for a History of Anthropological Theory*. Edited by Paul Erickson and Liam Murphy, 5th ed., University of Toronto Press, 2017, pp. 347-358.

Butler, J. *Gender Trouble: Feminism and the Subversion of Identity*, Routledge, 1999.

Butler, J. *Who's afraid of gender?* Farrar, Strauss, and Giroux, 2024.

Carnes, N. *Queer Community: Identities, Intimacies, and Ideology* (1st ed.). Routledge, 2019, https://doi.org/10.4324/9780429029424.

Fominaya, C. F. "Collective Identity in Social Movements: Central Concepts and Debates." *Sociology Compass*, 4: 393-404, 2010, https://doi.org/10.1111/j.1751-9020.2010.00287.x

Greene, L. *Drag queens and Beauty Queens: Contesting Femininity in the World's Playground*. Rutgers University Press, 2021.

HRC Staff. "Human Rights Campaign Applauds Governor Gretchen Whitmer for Establishing LGBTQ+ Commission, Continuing Historic Progress For LGBTQ+ Rights in Michigan." Human Rights Campaign, 12 June 2023, https://www.hrc.org/press-releases/human-rights-campaign-applauds-governor-gretchen-whitmer-for-establishing-lgbtq-commission-continuing-historic-progress-for-lgbtq-rights-in-michigan.

Kashiwagi, S. "Trump-appointed federal judge rules Tennessee's anti-drag show law is 'unconstitutional.'" CNN Politics, 3 June 2023, https://www.cnn.com/2023/06/03/politics/tennes-

see-drag-ban-law-unconstitutional/index.html.

Rasmussen, C., & Rasmussen, T. *Diary of a Drag Queen*. Farrar, Straus and Giroux, 2020.

Rothberg, Emma. "Sylvia Rivera." *National Women's History Museum*, 2021, https://www.womenshistory.org/education-resources/biographies/sylvia-rivera.

Rupp, L.J. & Taylor, V. *Drag Queens at the 801 Cabaret*. 2003. University of Chicago Press, 2015.

Shane, Cari. "The First Self-Proclaimed Drag Queen Was a Formerly Enslaved Man." *Smithsonian Magazine*, 9 June 2023, https://www.smithsonianmag.com/history/the-first-self-proclaimed-drag-queen-was-a-formerly-e nslaved-man-180982311/.

Taylor, V. Rupp, L.J. and Gamson, J. "Performing Protest: Drag Shows as Tactical Repertoire of the Gay and Lesbian Movement", Myers, D.J. and Cress, D.M. (Ed.) *Authority in Contention (research in Social Movements, Conflicts and Change, Vol. 25)*, Emerald Group Publishing Limited, Leeds, pp. 105-137, 2024, https://doi.org/10.1016/S0163-786X(04)25005-4.

Tucker, J. D. *Queens of Color: What It Means to Be a Black Drag Queen* (thesis), ProQuest, 2023.

David Adaway is a Computer Engineering major at the University of Memphis expecting to graduate in spring 2026. He is a member of the Helen Hardin Honors College, and he plans to pursue a master's degree after earning his bachelor's degree. He has been involved in research under the supervision of Dr. Chrysanthe Preza, Professor and Chair of the Electrical and Computer Engineering Department, since he joined her Vertically Integrated Projects team in fall 2022. He began using the Axio Imager.Z1 microscope in her lab since summer 2024.

Chrysanthe Preza is the Kanuri Professor and Chair of the Department of Electrical and Computer Engineering at the University of Memphis, where she joined in 2006. She received her D.Sc. degree in Electrical Engineering from Washington University in St. Louis in 1998. She leads the research in the Computational Imaging Research Laboratory. Since 2022, she is the founder and Director of the Vertically Integrated Projects (VIP) Program at the University of Memphis. Her research interests are imaging science, estimation theory, computational optical sensing and imaging applied to multidimensional multimodal light microscopy and hyperspectral imaging, and computational imaging enabled by deep learning. She received a CAREER award by the National Science Foundation in 2009, the Herff Outstanding Faculty Research Award in 2010 and 2015, and she was the recipient of the Ralph Faudree Professorship at the University of Memphis 2015-2018. Since 2018, she has been the recipient of the Ravi and Eleanor Kanuri Professorship. She was named Fellow of the SPIE in 2019 and Fellow of the Optica (OSA) in 2020. She served as Associate Editor for *IEEE Transactions on Computational Imaging*, as Topical Editor for Optica's *Applied Optics*, and as an Executive Editor for *Biological Imaging, Cambridge University Press.* She currently serves as an Editorial Board Member for *Journal of Imaging, Computational Imaging and Computational Photography section.*

Adaway and Preza's paper received a *QuaesitUM* outstanding paper award.

David Adaway & Chrysanthe Preza

Comparison of Two Three-Dimensional Imaging Approaches Used in Fluorescence Microscopy

Faculty Sponsor

Dr. Chrysanthe Preza

Abstract

This study introduces readers to fluorescence microscopy, the study of imaging biological samples that emit light when excited by light at specific wavelengths. A ZEISS Axio Imager.Z1 microscope was used to image an eosin-stained brine shrimp sample by *optical sectioning*, capturing the sample's full 3D depth in a widefield image stack. However, due to blurring caused by the microscope's point spread function (PSF), more complex imaging approaches must be used to obtain accurate images. This study focuses on two such approaches: deconvolution, applying a post-processing algorithm to remove the PSF from the widefield image; and structured illumination microscopy, during which a structured illumination pattern (provided by the ApoTome.2 add-on hardware) is projected on the sample during acquisition so more information is captured through light modulation, leading to fast and accurate image reconstruction. Readers who are unfamiliar with these microscopy modes can learn and apply them in their own research.

Introduction

Fluorescence microscopy is a subfield of microscopy often used when imaging biological samples [1]. These samples contain fluorophores: molecules that, when excited by certain wavelengths of light, will emit light at a different wavelength in return. A microscope can filter out the excitation wavelength so the resulting image only includes the emission wavelength; only materials that fluoresce will be visible, providing the greatest contrast between the labeled specimen and the background. Some samples (those that are autofluorescent) have natural light-emitting properties; other samples can be stained with fluorescent dyes for imaging purposes.

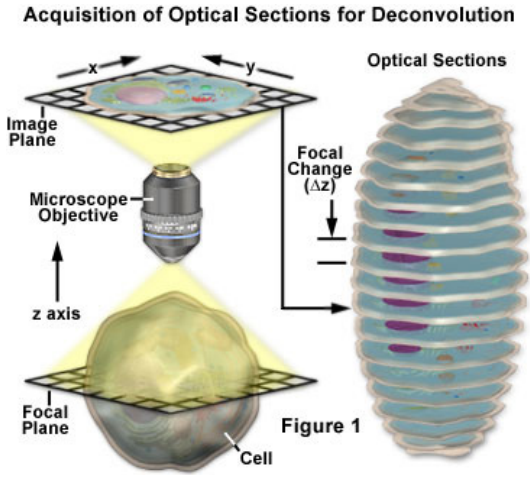


Figure 1.
3D image acquisition with uniform widefield microscopy [2].
One image of the subject is captured as a different part of the object is placed in the microscope's focal plane; the optical sections are stacked to form a 3D volume.

In a conventional light microscope, the sample is placed on a stage and illuminated by a lamp. For fluorescent samples, a filter is used to shift the light to the excitation wavelength. A two-dimensional (2D) image is then taken using a digital (CCD) camera integrated to the microscope, while the imaging lens is focused on the sample. This poses an issue for biological samples, many of which are three-dimensional (3D): in addition to the two lateral axes (X and Y), there is a third axis (Z) corresponding to the depth of the specimen, or the position within the specimen's thickness. One approach for imaging these samples is optical sectioning microscopy (OSM), where a 2D image is taken at each axial depth [3]. The depth is

changed by translating the stage vertically, which changes the part of the sample that is in focus by the microscope's objective, or imaging lens. The volume of the sample is represented by the "stack" of 2D images, or "slices" (Figure 1). For this method, the sample does not need to be specially prepared (besides the fluorescence staining, if necessary), and there is no risk of damage by physically sectioning it, unlike in electron microscopy for instance.

The drawback of optical sectioning microscopy is that each slice, also known as a widefield raw image, will contain both in-focus and out-of-focus light [3]. Emitted light from the entire volume is always visible to the objective, so regions of the sample that appear clear are mixed with regions that appear blurry; the position of the stage only changes which regions are in-focus. To obtain a clear depiction of a sample volume, the out-of-focus light must be removed from each slice.

The blurring from out-of-focus light can be modelled by a mathematical process called convolution, the integration of the product of the sample's emitted light and a shifted version of the microscope's point spread function (PSF), the microscope's response to a single point of light [4]. Each point of captured light is transformed according to the PSF (Figure 2); to remove out-of-focus light, the inverse process known as deconvolution must be performed. Widefield deconvolution microscopy uses the PSF to make educated guesses on the true appearance of the emitted light and solve the inverse imaging problem. There are different deconvolution algorithms of varying processing time and accuracy [5].

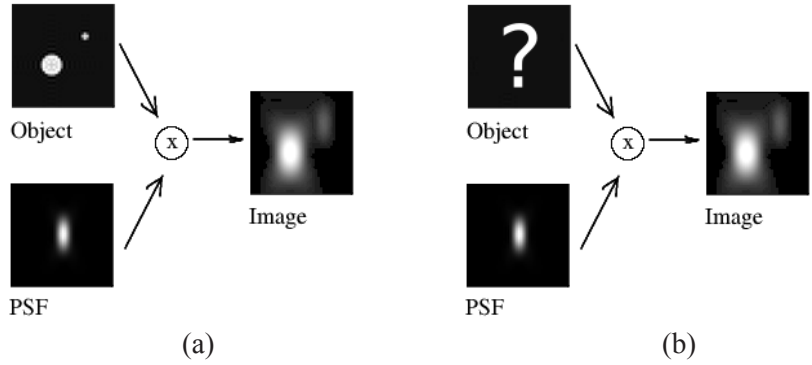
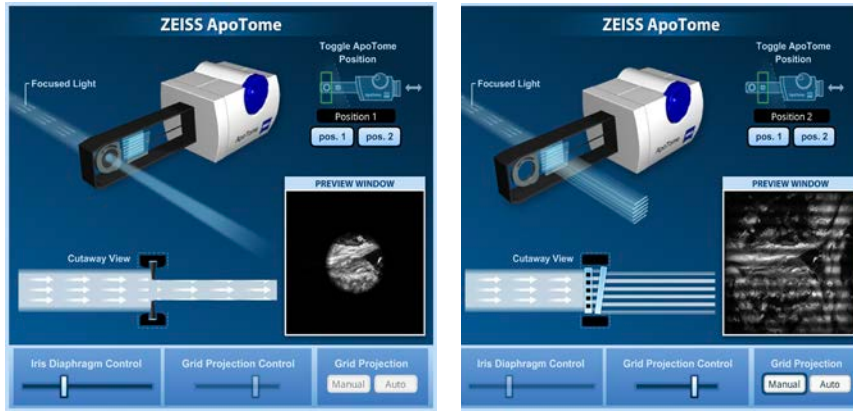


Figure 2.

(a) The output image is the convolution of the sampled object and the microscope's point spread function [6]. The PSF causes the light of the object to be redistributed accordingly. (b) In deconvolution, the image and the PSF are used to reconstruct the unknown object. The light transformation is reversed.

Alternatively, the optical sectioning process can be enhanced by structured illumination microscopy (SIM), removing the need for deconvolution. In SIM, a set of grid lines, or a structured illumination (SI) pattern, is applied on top of a 2D image (Figure 3). The SI pattern modulates the emitted light to different frequencies so that information that normally would not transfer through the widefield microscope can now be captured through this process. Each phase of the SI pattern (position of the grid lines) changes the modulation of light; thus, while summing the SIM images from different phases of the pattern results in a widefield image, image processing programs can extract information from each phase image and output a final result with improved optical sectioning [7]. The SIM process can be accomplished by taking multiple (usually, three to five) images at each axial depth of interest (the whole volume does not need to be processed) and using specialized microscope hardware to project a different SI pattern phase on each image. A computer program can then process the raw SIM data and output a new image at each chosen depth, improving the visual clarity of the image and the volume when a 3D image is desirable.

This study was performed by the primary author, an undergraduate student researcher, in order to learn about 3D fluorescence microscopy, image acquisition, computational image processing techniques, and data analysis. These are important preliminary skills necessary for future research inquiries. In addition, the work in this study was done in part to verify the correct operation of the equipment with a newer version of the commercial software, which was installed and tested for the experiments described below. As a demonstration of this work, this article will compare the processes and results obtained using conventional widefield deconvolution microscopy and optical sectioning structured illumination microscopy (OS-SIM, or simply "SIM" hereafter). First, the equipment used in this study and the actions taken to acquire the images are described. Then, various results from different processes are presented alongside technical explanations and visual observations. Finally, the images are directly compared (qualitatively and quantitatively) to determine the usefulness and application of each image restoration method.



(a) (b)

Figure 3.

The ZEISS ApoTome add-on device creates a structured illumination pattern that can be integrated in widefield microscopy [8]. (a) In position 1, the ApoTome is not engaged; the illumination is uniform resulting in conventional widefield microscopy. (b) In position 2, focused light is split into sections, which are projected onto the sample object. Different phases of the illumination pattern are created by adjusting the angle of the reflective glass inside the ApoTome, causing the lines superimposed on the image (here displayed in the "Preview Window") to shift up and down.

Methodology

This research study was performed in Dr. Chrysanthe Preza's Computational Imaging Research Laboratory (CIRL) in the Electrical and Computer Engineering Department at the University of Memphis. This is a computer-controlled microscopy experiment: specialized software is used to control the actions of the microscope. As part of this study, the undergraduate student contacted the microscope's manufacturer to obtain an updated version of the software; this required meetings with technical support to configure it correctly. Thanks to these preliminary steps, the microscope could be used for projects conducted in the CIRL. In this section, we will detail the actions taken to acquire, process, and prepare images using the tools available in our lab.

Tools and Parameters

All data acquisition and processing was done on a fixed sample of a brine shrimp stained with the fluorescent dye eosin. Image acquisition was performed using a motorized Axio Imager.Z1 microscope (Figure 4a), manufactured by ZEISS Microscopy. The images were captured using an Ax-

ioCam MRm camera, the one shown above the eyepieces in Figure 4a; captured images have a resolution of 1388 x 1040 pixels. The ApoTome.2 add-on device (or simply "ApoTome" hereafter, Figure 4b) projects the SI patterns used by OS-SIM [7]. Computer control of the microscope was enabled by the ZEN 3.8 software for Windows, which was also produced by ZEISS Microscopy. ZEN was used to control the acquisition and restoration parameters (Figure 5). The computer that ran ZEN had an Intel Xeon X5650 CPU (clock speed of 2.66 GHz) with integrated graphics; the lack of a dedicated GPU had a significant effect on computational time.

(a) (b)

Figure 4.
(a) The Axio Imager.Z1 microscope (modified with a side imaging path for another research project in the CIRL) used for this study. (b) The ApoTome.2 add-on hardware.

All image restoration methods were performed on a single Z-Stack acquisition of the specimen, done with the ApoTome in position 2 (Figure 3b). Three excitation filters were used: Rhodamine, EGFP, and DAPI. The images were captured using the Plan-APOCHROMAT 20x/0.8 objective dry lens: the magnification is 20 and the numerical aperture is 0.8. With this objective, each pixel is a square of side length 0.3225 µm, so the axial step size was set to 0.32 µm (ZEN does not allow more precise settings) so the stack would contain cubic voxels. 184 slices were captured, and each slice contains five phase images due to the ApoTome's SI pattern. The axial range was chosen such that the middle slice (#92) would contain the most in-focus data while the first and last slices would be the most out-of-focus. The entire depth of the specimen was traversed by remotely translating the microscope stage (where the brine shrimp sample was set) using the ZEN software.

In the Z-Stack acquisition, the axial movement of the stage, the changing of the excitation filters, and the adjustment of the SI pattern are performed automatically. The output of the acquisition is a "Raw data" image stack containing the phase images for each slice and excitation filter (2,760 total images). Afterwards, we used ZEN's "Conventional Fluorescence" option to obtain the widefield stack and the "Optical sectioning" option to obtain the SIM-reconstructed stack; ZEN performs the respective computations on the raw data stack. The widefield images were digitally processed by applying three of the built-in deconvolution algorithms: Nearest Neighbor [9], Regularized Inverse Filter (RIF), and Conjugate Gradient Maximum Likelihood (CGML, or "Constrained Iterative") [10].

For the RIF and CGML deconvolution methods, the microscope's 3D point spread function (PSF) is required; the algorithms use the PSF to determine how the light was transformed during the convolution process (Figure 2). ZEN can generate the PSF automatically using acquisition metadata contained in the widefield data file. Alternatively, an external PSF can be provided; this would be done to input PSFs obtained by empirical measurements. In this study, an auto-generated PSF was used for RIF, and a PSF provided by an experienced research scientist (who developed CGML [10]) was used for CGML. We note that Nearest Neighbor uses a 2D PSF generated by ZEN, and OS-SIM does not require any PSF.

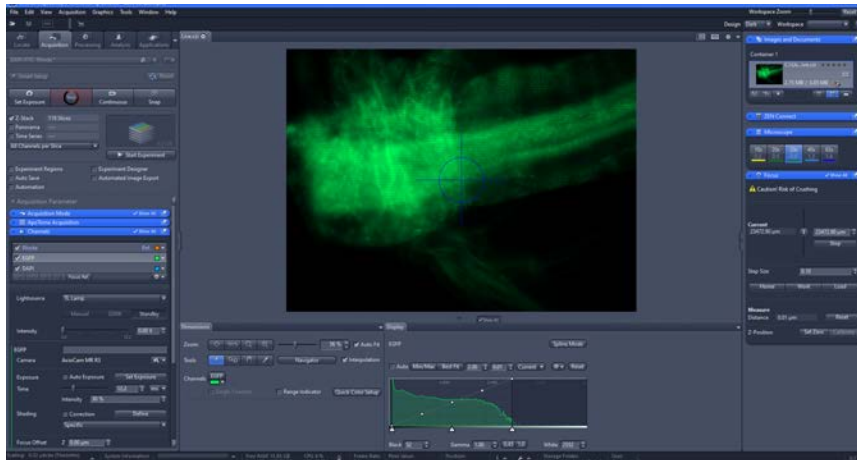


Figure 5.
The user interface of ZEN 3.8, the software used to control the acquisition. A live view of the microscope's camera is displayed. Illumination grid lines superimposed on the sample are evident in the image.

Point Spread Function Generation

For deconvolution, using an accurate point spread function (PSF) is important because the PSF acts as map for light between the sampled object and the image (Figure 2), much like a mathematical function f with input x and output y. In the same way it is impossible to solve the equation $y = f(x)$ for x without knowledge of f, the sampled object cannot be accurately reconstructed without knowledge of PSF. Acquiring the PSF is the most important step of a deconvolution!

After performing an acquisition, metadata such as the numerical aperture, the wavelength of light, and the lateral and axial dimensions of the images are stored in the CZI (Carl Zeiss Image) format alongside the images themselves, so ZEN is able to automatically generate a corresponding PSF [11]. ZEN provides two algorithms for calculating a PSF based on a scalar model and a vectorial model. In general, the vectorial model is more accurate [12], but more difficult to compute via numerical methods, which sometimes can lead to artifacts, as demonstrated in Figure 6.

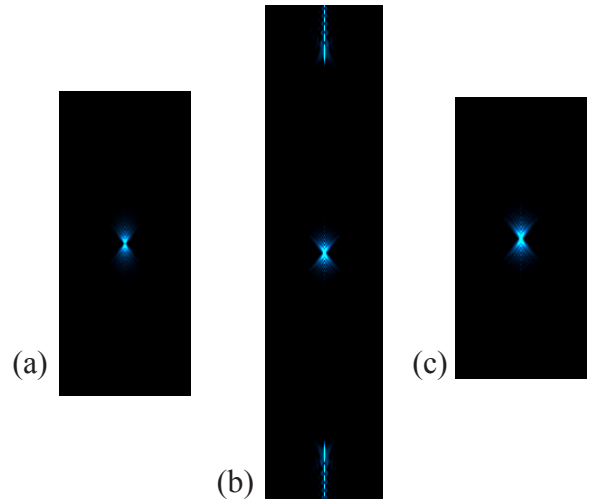


Figure 6.
Axial views of 3D PSFs (DAPI filter). (a) Scalar PSF for the widefield acquisition (Figure 7), generated by ZEN. (b) Vectorial PSF generated by ZEN. (c) Vectorial PSF supplied externally. Resolutions: (a) 109x240 pixels, (b) 109x454 pixels, (c) 109x218 pixels.

While the scalar PSF (Figure 6a) turned out aberration-free, the vectorial PSF (Figure 6b) has visible "tail-artifacts" at the top and bottom of the images. To obtain the highest-quality deconvolution, an expert re-

search scientist (co-developer of the CGML method [10]) was consulted; he provided a new vectorial PSF (Figure 6c), which was used to perform the CGML method. For the Regularized Inverse Filter deconvolution, we used the generated scalar PSF, while for the Nearest Neighbor a 2D PSF was used as noted above.

Figure Preparation

All results from ZEN were exported in the CZI format and imported into Fiji [13] (a distribution of the open-source imaging software ImageJ [14]) using its Bio-Formats Importer plugin [15]. Fiji was used for preparation of the figures presented here for image comparison and analysis. Each dataset (widefield, deconvolution results, and OS-SIM restoration) is an image stack containing 184 slices (traversing 58.56 microns of the sample's depth). Each slice has a resolution of 1388x1040 pixels (or 447.63x335.40 microns). Figures of the XY view in this paper depict slice #92, the middle slice. The XZ orthogonal views allow the sample's depth to be visualized: the x-axis increases rightward, and the z-axis increases downward. The XZ views are focused on Y = 520, the horizontal center of the XY view.

This study includes a qualitative and quantitative analysis and comparison of results from different image restoration methods. By examining intensity profiles and orthogonal views, each method's improvement over widefield microscopy is demonstrated.

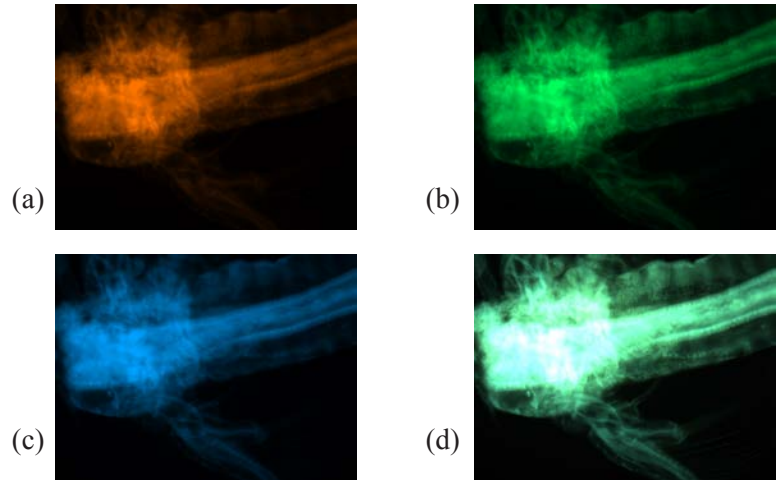


Figure 7.
XY views of acquired widefield image stacks, using the (a) Rhodamine, (b) EGFP, and (c) DAPI excitation filters. (d) Shows the composite image created from (a)-(c).

Results

In this section, we explain the computational methodologies involved in the creation of each image. The results are compared by examining the images visually (qualitatively) and by examining their pixel intensities (quantitatively).

Conventional Fluorescence

The conventional fluorescence, or widefield, dataset studied here was obtained by summing the phase images from the ApoTome optical sectioning acquisition. There is one 2D image ("slice") for each axial depth; the slices form a stack representing the sample volume, as in Figure 1. Figure 7 shows the middle slices from four widefield image stacks from the brine shrimp sample; three stacks were acquired by using different excitation filters to shift the illumination wavelength, and the fourth stack contains composite images created by overlaying the images from the first three stacks.

Each excitation filter exposes different regions of the sample; the part that gets excited and consequently reflects light at a different wavelength. The images appear somewhat blurry; all the image restoration methods attempt to improve the clarity of these images by removing noise and adjusting out-of-focus light so more of it is in focus.

When the sample was acquired, care was taken to not introduce any errors into the output images. However, the range of depths acquired was too small, resulting in some light not being captured; this is clear from the XZ view (Figure 8a), as the light extends past the top and bottom of the view. This error had effects on the point spread function (Figure 6), and therefore, the deconvolutions.

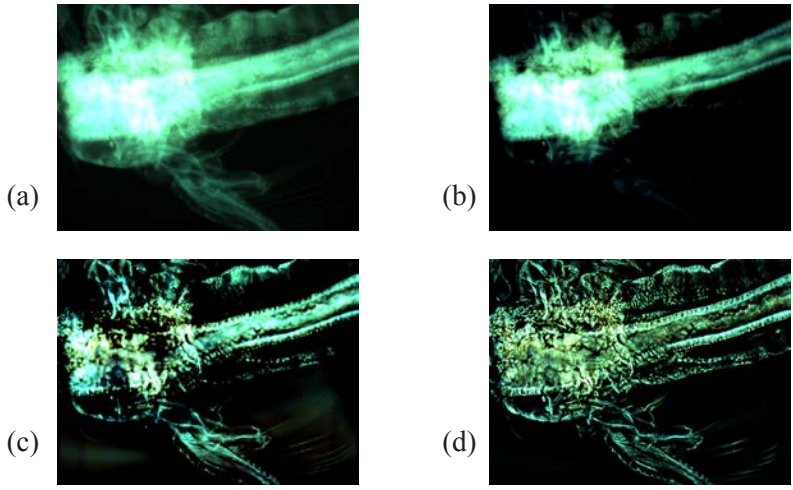


Figure 8.
XY views of the (a) raw widefield image, and restored images obtained using: (b) Nearest Neighbor, (c) Regularized Inverse Filter, and (d) CGML deconvolutions.

Widefield Deconvolution

Three deconvolution algorithms were applied to the raw widefield images. These are postprocessing methods that generate and attempt to remove the microscope's point spread function blurring effect from the acquired widefield data. Figure 8 and Figure 9, respectively, present the XY and XZ views from widefield and each algorithm for comparison. These images are composites of the three excitation filters used in the widefield acquisition.

Nearest Neighbor (Figure 8b) is the oldest and fastest 2D deconvolution algorithm provided in the ZEN software. It assumes that for each plane, the two immediate adjacent planes contribute the most information [3]. By removing convolved light from the nearest neighboring planes, an approximation of the plane of interest can be obtained. Only two 2D convolution operations need to be processed for each plane.

The Nearest Neighbor 2D deconvolution was performed using ZEN's default parameters. Compared to the widefield image (Figure 8a), there are less regions visible, but the regions that are visible appear mostly the same.

Regularized Inverse Filter (RIF, Figure 8c) is the next-fastest deconvolution algorithm in ZEN [16]. Unlike Nearest Neighbor, it performs a 3D deconvolution, and it attempts to mitigate the effects of noise on the

inverse image (the widefield output) by applying a denoising filter (regularization) before deconvolution.

The RIF deconvolution was performed using ZEN's default parameters, including the Zero Order regularization model and an automatically set strength. Comparing their orthogonal views of the XZ axis (Figure 9), the RIF deconvolution was able to restore more in-focus light than Nearest Neighbor.

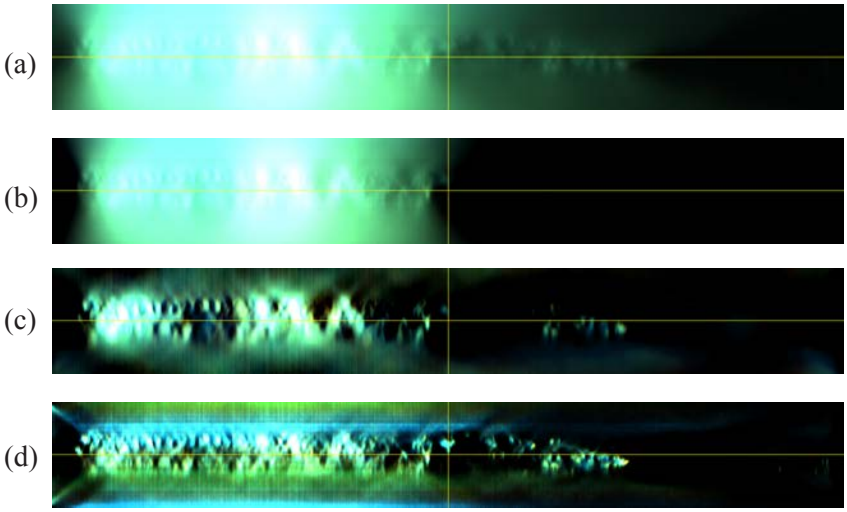


Figure 9.
XZ views of the (a) widefield image, and restored images obtained using: (b) Nearest Neighbor, (c) Regularized Inverse Filter, and (d) CGML deconvolutions.

The Conjugate Gradient Maximum Likelihood (CGML, Figure 8d) algorithm, unlike the others presented here, is an iterative method; the same steps are repeated multiple times until a satisfactory result is obtained based on a metric. In general, in each iteration, an estimate of the sampled object is convolved with the PSF and compared to the widefield image [5]. The percent error, the difference between the convolved estimate and widefield images, is used to generate a new estimation of the sample with reduced error. To aid in estimation, a non-negativity constraint is applied on the estimation, which means no pixel can have an intensity below zero, consistent with the fluorescence intensity in the sample.

This CGML deconvolution was performed with a maximum of 100 iterations per excitation filter, though only the result from the DAPI filter

data used that many. For the result from the Rhodamine data, 78 iterations were used, and for the EGFP result, 92. The algorithm is designed to stop when the percent error is satisfactory, or it cannot be decreased further. Comparing the images in Figure 8, it is clear that the CGML result (Figure 8d) contains the most detail.

The different results are also compared by their XZ views (Figure 9). These orthogonal images allow the depths (z-axis) of each image stack to be visualized. These XZ views are focused on Y = 520, which is the horizontal center of each XY view. X increases from left to right, and Z increases from top to bottom. The improvements to the XZ view mirror the improvements to each XY view. Nearest Neighbor (Figure 9b) restores some light but otherwise looks similar to widefield (Figure 9a). RIF (Figure 9c) improves upon Nearest Neighbor by removing much of the blur and returning light to the in-focus region. In CGML (Figure 9d), more in-focus light is visible, though artifacts can be seen in the out-of-focus regions (the top and bottom of the image).

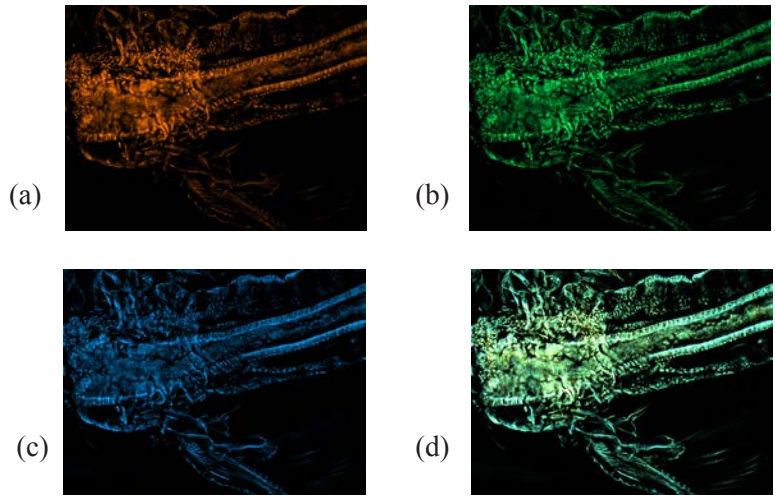


Figure 10.
XY view of the CGML deconvolution. (a) Rhodamine, (b) EGFP, (c) DAPI excitation filters, and (d) composite image of the three filters.

Figure 10 is a breakdown of the CGML image in Figure 8d, showing the result of each excitation filter alongside the composite image. The improvements observed in the composite image are present in each channel (compare to the widefield images in Figure 7).

Structured Illumination Microscopy

Structured illumination microscopy (SIM) datasets are reconstructions of an object captured using the optical sectioning capabilities of the ApoTome.2 add-on hardware. Optical sectioning is enabled by projecting a structured illumination (SI) pattern onto the captured images. The SI pattern modulates the illumination from the microscope's lamp so that the captured image will have light and dark sections corresponding to the structured light (Figure 3b); the light portions modulate (multiply) the in-focus information present in the slice [7]. By changing the lateral position (phase) of the pattern, the SI pattern multiples different portions of the image that are in focus.

In this study, the entire volume of the sample was imaged, and five phase images were captured for each slice. Each phase image contains a shifted version of the same sinusoidal grid pattern overlay. The different phase images are subtracted from each other, then the differences are squared and summed. Finally, the square root of the sum is taken, resulting in an image in which the grid pattern is removed along with the out-of-focus light, creating a new SIM-reconstructed image containing only the optically sectioned (in-focus) portions of each phase image [7].

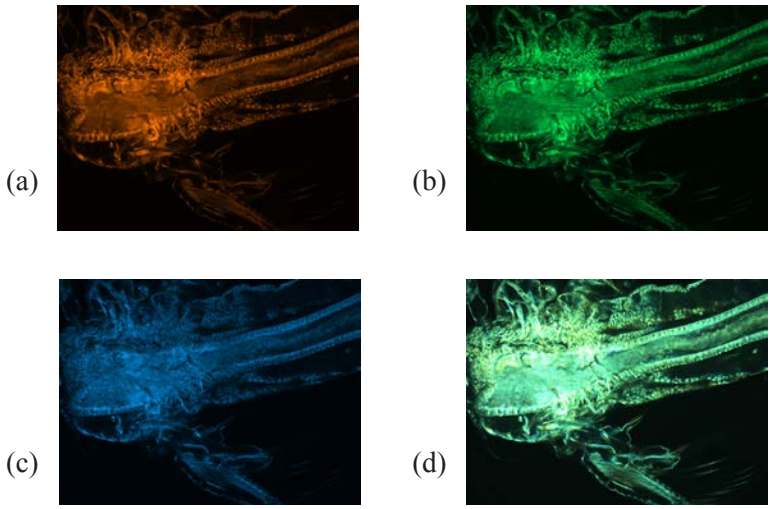


Figure 11.
XY views from the acquired and processed optically sectioned SIM image stack for: (a) Rhodamine, (b) EGFP, (c) DAPI excitation filters, and (d) composite image of the three filters.

ZEN's algorithm also applies corrections to prevent errors from the acquisition (unstable illumination, incorrect pattern positions, bleaching, etc.) from affecting the reconstruction as much as in deconvolution [17]. The effects of this computational approach of ApoTome SIM on the sample data can be seen in Figure 11.

Not only are the SIM images (Figure 11) less blurred than the widefield images (Figure 7), but they also display similar details as in the CGML images (Figure 10). The head (center left) of the sample and the whiskers (bottom middle) are much clearer in the SIM reconstruction than the widefield images, yet the result is still different from CGML due to the different processes involved. The next section contains more detailed analyses of the widefield, CGML, and SIM images.

Qualitative Comparison

Using the tools featured in Fiji [13], this study compares the widefield, CGML, and SIM images to determine each image restoration method's advantages and disadvantages. Figure 12 shows the XY views of the middle slice from each stack.

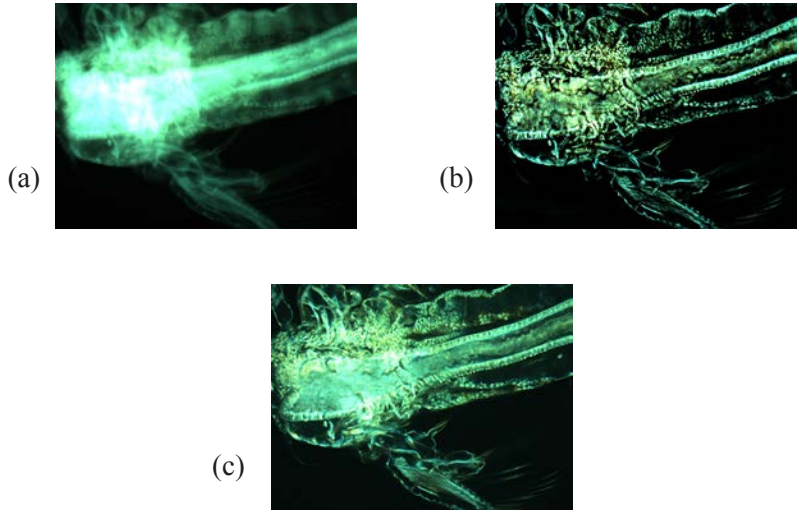


Figure 12.
XY views from the (a) widefield, (b) CGML, and (c) SIM image stacks (composite image of the three excitation filters).

Compared to the widefield slice, the CGML and SIM slices display the brine shrimp's structures in much greater detail. In particular, the head of the shrimp (left center) was completely blurred in the widefield slice; both restorations are sharper in that area. The CGML and SIM slices provided similar, yet slightly different views of the subject. CGML appears to contain more contrast in the structures, but without knowledge of the sample's true appearance, neither restoration can be objectively ruled as more accurate. Differences can also be seen in the XZ views, shown in Figure 13.

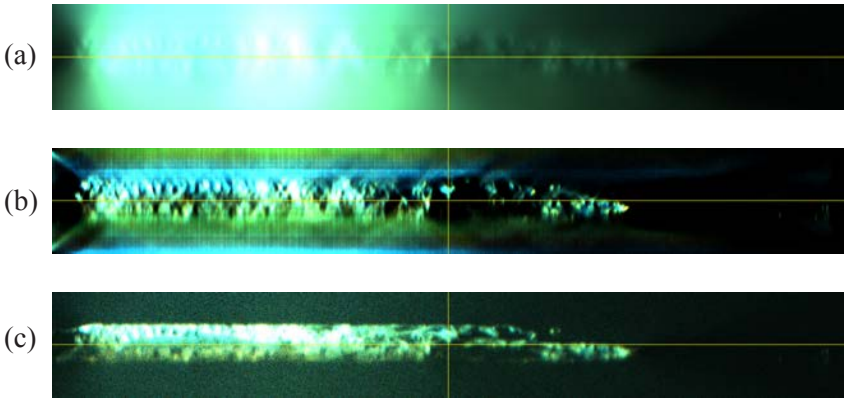


Figure 13.
XZ views from the (a) widefield, (b) CGML, and (c) SIM image stacks (composite image of the three excitation filters).

For each orthogonal view, in-focus light appears near the center of the sample (marked by the yellow horizontal line). In the widefield view, there is a lot of light above and below the in-focus region. In the CGML view, there is less blurring and more detail in the in-focus region, and the light in the out-of-focus region is also less blurred, though still present due to some artifacts. The SIM view appears to smooth out the in-focus region by having more light in the region but also lower contrast in the structures. The out-of-focus light is less prominent, with only some slight bleaching present. There is a progression in image quality from raw widefield to processed widefield, with SIM having a mixture of positive and negative differences.

Quantitative Comparison

Changes in pixel data within a certain region of the image can be visualized by intensity line plots, which plot a gray pixel value (of an arbitrary unit) against distance within the image of the sample. The lowest possible

intensity value (0) appears black, and the highest possible value appears white. Each image restoration method has a different result, so an intensity plot from the same region will look different for each case.

In this study, the intensity plots were created in Fiji [13]. A yellow line is drawn over the measured pixels, and the graph plots each pixel's intensity (referred to as "Gray Value" in Fiji) against the distance in microns from the left endpoint of the line. For easier viewing of intensities, images are presented in grayscale. Figure 14 shows an intensity plot from a widefield image.

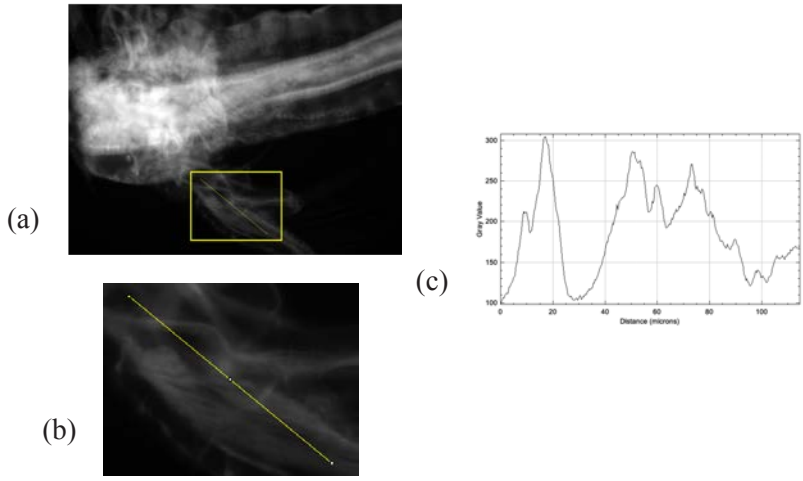


Figure 14.
(a) XY view from the widefield image stack (DAPI filter). (b) Crop from (a) shown by the yellow rectangle, focused on the 115 micron-long line overlay. (c) Intensity plot across the line overlay.

Figure 14c plots pixel intensity values against the distance of the line in Figure 14b. The widefield image uses 12 bits to store pixel data, so the maximum possible intensity value is $2^{12} - 1$, or 4,095. The pixels captured in this line overlay have values from 100 to 300. The plot contains smooth curves, which is expected due to the blurring effects of the PSF of the biological sample features. It appears that many of the sample's details have been lost, and out-of-focus information (blur) creates smooth curves. Figure 15 shows a different intensity plot from the same region of the result obtained after processing with the CGML deconvolution method.

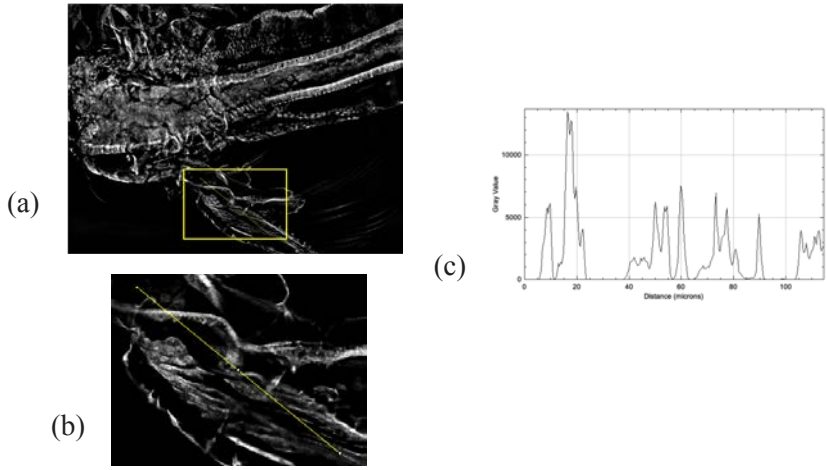


Figure 15.
(a) XY view from the CGML image stack (DAPI filter). (b) Crop of (a) shown by the yellow rectangle, focused on the 115 micron-long line overlay. (c) Intensity plot across the line overlay.

The results from the CGML deconvolution are 16-bit images, so the maximum intensity is $2^{16} - 1$, or 65,535. Much of the line selection shows an intensity of 0 (due to the algorithm's non-negativity constraint), but the peaks are better resolved (more peaks are visible, and they have defined maximums) and have values from roughly 5000 to 15000, which are higher compared to the widefield image. The deconvolution restores visual data (the relative brightness of each pixel). The peaks and curves are expected to be closer to the actual details in the sample, given the increased contrast and resolution in the image. Figure 16 shows a similar plot from the SIM image restoration.

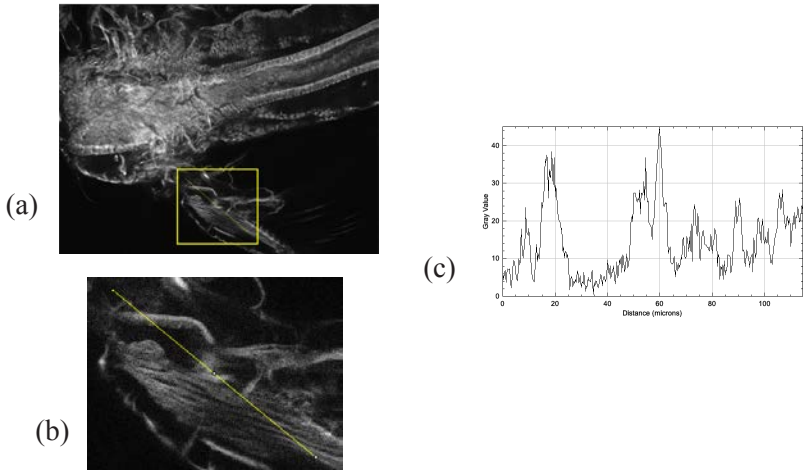


Figure 16.
(a) XY view from the SIM image stack (DAPI filter). (b) Crop of (a) shown by the yellow rectangle, focused on the 115 micron-long line overlay. (c) Intensity plot across the line overlay.

The SIM image, like the widefield image, is 12 bits, so the maximum intensity is 4,095, although the largest plotted intensity is much lower because SIM processing involves subtraction operations as part of the processing method [7]. As a result, the numerical data is lost, and the images appear more grainy or noisy than the CGML result. However, the pattern in this intensity plot is closer to the one in the CGML plot than the widefield plot, as the main peaks appear in the same locations. However, pixels that were zeroed out in CGML have values here, so more data is present in SIM, which could still be residual blur, as fewer peaks are identified. Additionally, there is more variation in the value of adjacent pixels, which better matches expectations when data is noisy. Although SIM and CGML XY images look similar, it appears that the CGML deconvolution provides an image with higher contrast, more resolution, and less noise than the SIM image.

Discussion

Overall, both widefield deconvolution microscopy (through the CGML algorithm) and optically sectioned structured illumination microscopy (OS-SIM, powered by ApoTome.2) were effective methods to restore an image of a fluorescent biological sample. CGML and SIM had similar results when applied to the experimental widefield image, with CGML having the greater contrast. However, OS-SIM outperforms CGML in process-

ing time and ease of use. The ApoTome acquisition and the simpler SIM computations finished in minutes instead of the several hours necessary for the iterative computations of the CGML method, although both processes would have been faster with a dedicated graphics card. Since no point spread function is required, OS-SIM is simpler for a layperson to operate; by contrast, deconvolution has many more parameters to adjust. Also, SIM is a per-slice operation, so a selected number of slices can be optically sectioned and processed. Deconvolution requires a large z-axis to be effective, which is apparent by the artifacts in the reconstructions (Figure 9). However, deconvolution does not require any additional hardware, and many open-source software implementations are available, so it is a much less expensive alternative restoration method to ApoTome SIM for experienced researchers. Otherwise, some training is required.

Conclusion

Three products from ZEISS Microscopy – the Axio Imager.Z1 microscope, the ApoTome.2 add-on hardware, and the ZEN imaging software – were used to acquire images from a fluorescent biological sample and perform several image restorations to make the images appear closer to the underlying object. No definitive conclusion could be made regarding the best restoration method since the true appearance of the sample is unknown; however, both SIM and CGML performed acceptably, and each could be used for different circumstances.

Most importantly, this study demonstrates two powerful techniques that have been developed over the last three decades to image 3D fluorescent samples and have been applied successfully to fields ranging from biology to health sciences. The two techniques appear in the literature for research studies that rely on close examination of fluorescent (naturally or otherwise) subjects, such as microorganisms or plant and animal cells. The brine shrimp sample here was used for training and demonstration purposes. The latest hardware and software are not necessary: the Axio Imager.Z1 and ApoTome.2 have been discontinued, and ZEN 3.8 is outdated, yet these tools worked here. Researchers that aim to use these products do not need to be experts in microscopy, either; however, they need to receive training and learn how deconvolution works. This study was performed by an undergraduate student who hopes to be able to use this knowledge to image samples for a collaborative biological study in the future.

Acknowledgement

I, David Adaway, would like to thank the National Science Foundation (NSF) for funding my research as part of the VIP Program (NSF-IUSE award #2120819) as well as the Helen Hardin Honors College for providing a fellowship grant allowing me to work during the summer of 2024. I am grateful to Dr. Chrysanthe Preza for her supervision in conducting the work and writing this paper, and for the training she provided so that I can use her research facilities, where I conducted this study. I would like to thank Carl Zeiss Microscopy, Inc, for providing me with a demo dongle for ZEN so that I could conduct this experiment. I would also like to thank Mr. Lutz Schaefer of AIMC, Inc, and consultant for ZEISS Microscopy, who has authored several of the cited works ([2], [5], [10], [17]) and has provided accurate vectorial PSFs, which allowed me to obtain a high-quality CGML deconvolution result. His expertise and useful discussions were instrumental in completing this study successfully.

References

[1] J. W. Lichtman and J.-A. Conchello, “Fluorescence microscopy,” *Nat. Methods*, vol. 2, no. 12, pp. 910–919, Dec. 2005, doi: 10.1038/nmeth817.

[2] W. Wallace, L. H. Schaefer, and J. R. Swedlow, “Digital Image Processing - Introduction to Deconvolution,” Olympus Life Science. Accessed: Jan. 04, 2025. [Online]. Available: https://www.olympus-lifescience.com/en/microscope-resource/primer/digitalimaging/deconv olution/deconintro/

[3] D. A. Agard, “Optical Sectioning Microscopy: Cellular Architecture in Three Dimensions,” *Annu. Rev. Biophys. Bioeng.*, vol. 13, no. 1, pp. 191–219, Jun. 1984, doi: 10.1146/annurev.bb.13.060184.001203.

[4] J.-B. Sibarita, “Deconvolution Microscopy,” in *Microscopy Techniques*, J. Rietdorf, Ed., Berlin, Heidelberg: Springer, 2005, pp. 201–243. doi: 10.1007/b102215.

[5] W. Wallace, L. H. Schaefer, and J. R. Swedlow, “A Workingperson’s Guide to Deconvolution in Light Microscopy,” *BioTechniques*, vol. 31, no. 5, pp. 1076–1097, Nov. 2001, doi: 10.2144/01315bi01.

[6] “Point Spread Function (PSF),” Scientific Volume Imaging. Accessed: Jan. 12, 2025. [Online]. Available: https://svi.nl/Point-Spread-Function-%28PSF%29

[7] M. A. A. Neil, R. Juškaitis, and T. Wilson, “Method of obtaining optical sectioning by using structured light in a conventional microscope,” *Opt. Lett.*, vol. 22, no. 24, p. 1905, Dec. 1997, doi: 10.1364/OL.22.001905.

[8] T. B. Gines and M. W. Davidson, “Structured Illumination: ZEISS ApoTome,” ZEISS Microscopy Online Campus. Accessed: Jan. 12, 2025. [Online]. Available: https://zeiss-campus.magnet.fsu.edu/tutorials/opticalsectioning/apotome/indexflash.html

[9] K. R. Castleman, *Digital image processing*. in Prentice-Hall signal processing series. Englewood Cliffs, N.J: Prentice-Hall, 1979.

[10] L. H. Schaefer, D. Schuster, and H. Herz, “Generalized approach for accelerated maximum likelihood based image restoration ap-

plied to three-dimensional fluorescence microscopy," *J. Microsc.*, vol. 204, no. 2, pp. 99–107, Nov. 2001, doi: 10.1046/j.1365-2818.2001.00949.x.

[11] S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *J. Opt. Soc. Am.A*, vol. 9, no. 1, p. 154, Jan. 1992, doi: 10.1364/JOSAA.9.000154.

[12] J. Kim, T. Li, Y. Wang, and X. Zhang, "Vectorial point spread function and optical transfer function in oblique plane imaging," *Opt. Express*, vol. 22, no. 9, pp. 11140–11151, May 2014, doi: 10.1364/OE.22.011140.

[13] J. Schindelin *et al.*, "Fiji: an open-source platform for biological-image analysis," *Nat. Methods*, vol. 9, no. 7, pp. 676–682, Jul. 2012, doi: 10.1038/nmeth.2019.

[14] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nat. Methods*, vol. 9, no. 7, pp. 671–675, Jul. 2012, doi: 10.1038/nmeth.2089.

[15] M. Linkert *et al.*, "Metadata matters: access to image data in the real world," *J. Cell Biol.*, vol. 189, no. 5, pp. 777–782, May 2010, doi: 10.1083/jcb.201004104.

[16] W. Wang and M. K. Ng, "Convex regularized inverse filtering methods for blind image deconvolution," *Signal Image Video Process.*, vol. 10, no. 7, pp. 1353–1360, Oct. 2016, doi: 10.1007/s11760-016-0924-3.

[17] L. H. Schaefer, D. Schuster, and J. Schaffer, "Structured illumination microscopy: artefact analysis and reduction utilizing a parameter optimization approach," *J. Microsc.*, vol. 216, no. 2, pp. 165–174, Nov. 2004, doi: 10.1111/j.0022-2720.2004.01411.x.

Farraday Johnson earned her B.S. in Biomedical Engineering with a minor in Chemistry from the University of Memphis in the Fall of 2024, where she was an Emerging Leader Scholar and Urban STEM Scholar. She began research as a Helen Hardin Honors College Summer Fellow, studying dual drug delivery, and later participated in the UTHSC and UPenn STAR research programs, focusing on Alzheimer's Disease and Related Dementias. As an undergraduate, she conducted ultrasound research in Dr. Carl Herickhoff's lab, presenting at the International Symposium for Ultrasonic Imaging and Tissue Characterization and earning 1st place at the University of Memphis's Student Research Forum. Now a Clinical Research Coordinator at the University of Pennsylvania, Farraday aspires to become a neurologist and is passionate about advancing Alzheimer's disease research.

Carl Herickhoff was born in Mankato, MN, in 1982. He earned a B.S. in Electrical Engineering from the University of Notre Dame in 2005, and a Ph.D. in Biomedical Engineering from Duke University in 2011. His professional career in ultrasound devices and systems research has included time at GE Global Research Center, Philips, Healthcare, Duke University, and Stanford University, before joining the Biomedical Engineering Faculty at the University of Memphis in 2020. Dr. Herickhoff received the NSF CAREER Award in 2023. His current research interests include transcranial ultrasound blood flow detection, super-resolution functional ultrasound imaging, breast cancer ultrasound screening, and quantitative ultrasound computed tomography techniques.

Farraday Johnson & Carl Herickhoff

Simulation of a 2D Array Transducer Design for Transcranial Doppler Signal Acquisition

Faculty Sponsor

Dr. Carl Herickhoff

Abstract

Timely and accurate diagnosis of strokes is critical for effective treatment to prevent further neuronal cell death and functional impairment. Acquisition of transcranial doppler (TCD) ultrasound signals to detect stroke is challenging due to the need to find an acceptable acoustic window through the skull and intersect the beam with the middle cerebral artery. We investigated the design of a novel 2D array transducer intended to interrogate skull thickness and allow steerability of the ultrasound beam to facilitate TCD signal acquisition. Field II, a MATLAB program that simulates ultrasound transducer fields, was used to model multiple 2D array geometries using triangular transducer elements inscribed within a 2 cm diameter circle. The results obtained suggest that a simplified 2D array transducer could be fabricated to provide TCD users with feedback to aid in optimal signal acquisition.

Introduction

Each year, there are 800,000 strokes in the United States leading to an average of 150,000 deaths, at a cost of $50 billion.[1] A stroke is a disruption of blood flow to the brain, and there are two types of strokes: hemorrhagic and ischemic. Hemorrhagic stroke occurs when there is a leakage of blood to surrounding areas within the brain due to a rupture of a cerebral vessel. Ischemic stroke occurs when there is a blockage of an artery within the brain by a blood clot or vascular plaque. Ischemic strokes account for 87% of strokes, with large-vessel occlusions (LVOs) as the most severe ischemic strokes; LVOs occur in 38% of ischemic stroke cases and account for over 90% of stroke deaths.[1,2]

The accurate identification of stroke type is critical because the treatment for each stroke type would exacerbate the other type. Furthermore, delay in identifying and treating stroke in a timely manner can lead to disability (e.g., physical or cognitive impairment) or death. For instance, ischemic stroke can be treated with tissue plasminogen activator (tPA, a blood thinner), and must be administered within 4.5 hours of onset of the stroke.[3] If a hemorrhagic stroke were to be treated using tPA, bleeding within the surrounding areas of the brain would be exacerbated, possibly leading to further neurologic issues. Currently, stroke type is determined (and LVOs are identified) by computed tomography angiography (CTA).[4]

Transcranial Doppler

Transcranial doppler ultrasonography (TCD) is currently used as an inexpensive way to monitor stroke patients by measuring the blood flow velocity in cerebral arteries.[5] Unlike CTA, TCD does not involve radiopaque contrast dye or ionizing radiation. Current TCD devices consist of a circular single-element transducer that generates a fixed beam, so the user must move the device manually to search for overlap of the beam with a major artery, such as the middle cerebral artery (MCA), to obtain a blood flow signal. To our knowledge, no clinical TCD devices incorporate a non-imaging transducer array. The other challenge for clinical users of this device is locating a sufficient acoustic window—the thinnest region of skull bone that allows optimal ultrasound beam transmission.

Thus, current TCD devices require optimization of four degrees of freedom (translation and rotation about x and y over the temporal region of the head) by trial and error. The goal of this project is to design a simple 2D array to both (1) sample the skull thickness underlying the array area and return a 2D profile to help locate an acoustic window, and (2) elec-

tronically steer the ultrasound receive beam in multiple directions off-axis and return a suggestion of angular tilt to optimize beam-vessel alignment and maximize blood flow signal. This array device would enable unskilled users to quickly and confidently detect the presence or absence of MCA blood flow and rapidly determine when a patient has an LVO stroke.

Methods

6-element array

Unlike current circular single-element or linear-array transducer geometries, a transducer array was fabricated using triangular elements to aid in beam optimization and steering. Utilizing the MATLAB-based Field II ultrasound simulation package, a novel 2D array of triangular elements was designed in **Figure 1**, using six elements inscribed in a 2 cm circle, the approximate diameter of current TCD devices. The sides of the inner triangular elements measured 10 mm each. The kerf, or distance, between each element was 0.04 mm. Beam patterns in the X-Z and Y-Z planes were simulated at focal ranges at 1 cm and 6 cm.

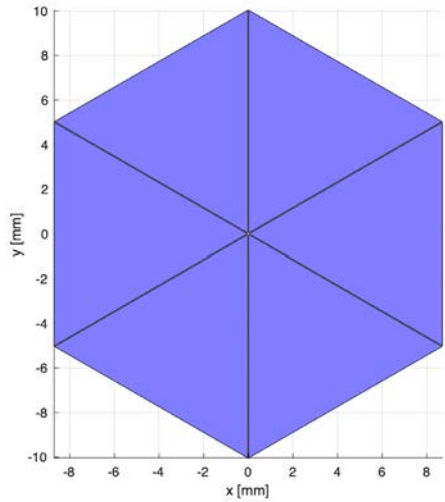


Figure 1.
2D element array consisting of 6 triangular elements

12-element Array

In Figure 2, a 12-element array was similarly configured and inscribed within a 2 cm circle. The sides of each element measured at 5.73 mm. The kerf was set to 0.04 mm. Beam patterns in the X-Z, Y-Z, and X-Y planes were simulated. For the X-Y plane, Z was initially set to 5 mm and then adjusted to the beam maximum. Focal ranges of [1:1:6] cm were simulated. Next, the array apodization (element-by-element weighting) was manipulated so that the response under individual elements was simulated. Apodization was constructed by activating and deactivating select transducer elements to minimize side lobes and modify the beam profile. This addresses the long-term goal of sampling and profiling the skull thickness. Finally, the beam for the whole array was steered off-axis by 10° in the azimuth (X) direction.

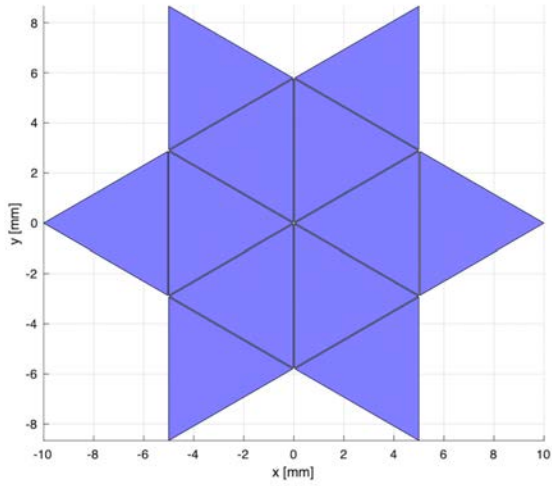


Figure 2.
Novel 2D element array consisting of 12 triangular elements.

Results and Discussion

6-element Array: Focal Range Control for X-Z and Y-Z Planes

In **Figure 3**, the X-Z and Y-Z planes are focused at 1 cm and 6 cm. The 6 cm focal range has a deeper focus; however, as shown in **Figure 3,** there was no observable difference in the beam patterns. This is due to the absence of curvature in the delay profile (only two transducer elements' transmit

timing cannot be phased to generate a concave wavefront converging to a focus). Since only two source points are symmetrically positioned about the origin, the resulting wavefront does not have the curvature necessary for beam focusing. The same occurrence was observed in the orthogonal plane. In summary, a curved delay profile cannot be achieved when an array provides only two elements along a given line through the origin.

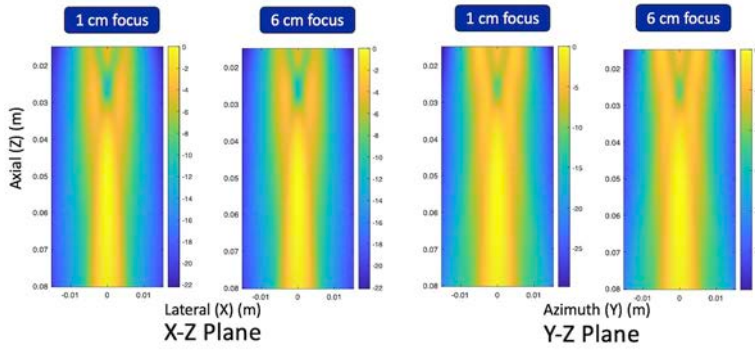


Figure 3.
X-Z and Y-Z plane simulation of beam patterns of the 6-element array at different focal depths. Color bar units are in decibels with respect to the maximum.

12-element Array: Focal Range Control for X-Z Plane and Y-Z Planes

In contrast to the 6-element array's inability to control its focal distance, the 12-element array increases the number of elements along a line through the origin to four. Due to the increase in elements in a line through the origin, a delay profile can be implemented to exhibit a concave curvature, which enhances beam focusing. In the graph, note the black dots that indicate the points of highest beam intensity. These points illustrate how the focal depth progresses as the distance increases. In **Figure 4**, when the focus is set to 1 cm, the beam is focused at 1 cm as expected. However, as the focal distance progresses beyond 35 mm, it is observed that the beam is unable to focus past this approximate far-field transition point defined by Fourier Optics (D^2/4*wavelength).

In the Y-Z Plane in **Figure 4**, the same behavior as stated for the X-Z plane is seen in the Y-Z plane. The beam's focus can be controlled to 35 mm.

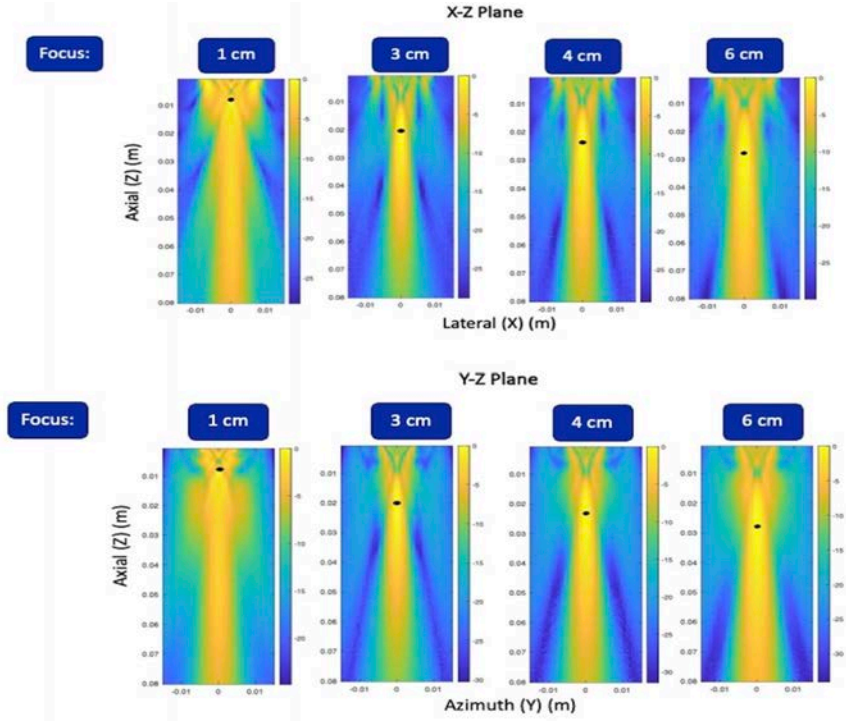


Figure 4.
X-Z Plane (top) and Y-Z Plane (bottom) simulation of beam patterns of 12 element array at varying focal depths. Color bar units are in decibels with respect to the maximum.

12-element Array: X-Y Planes at Beam Maximum

Additionally, in **Figure 5**, simulations were run to determine the maximum of the beam that could be focused. In the figures below, the X-Z plane and X-Y planes are displayed with the beam pattern perpendicularly sliced in the X-Y plane. The slices were taken in the X-Y plane at 5 mm and 35 mm.

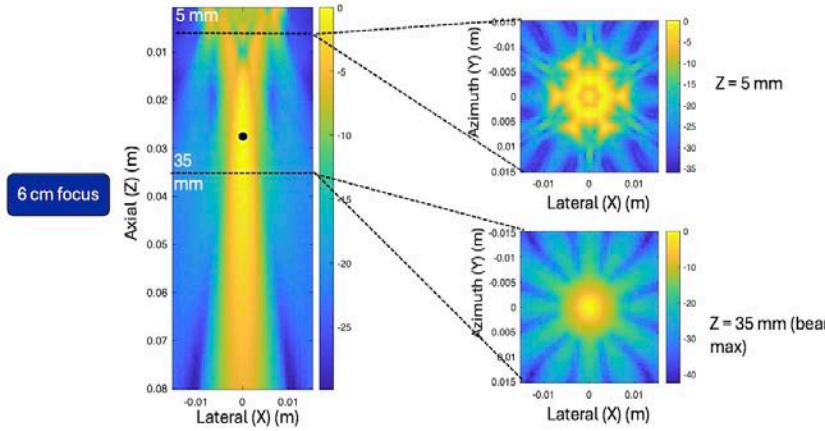


Figure 5.
The X-Z plane and X-Y planes are shown with a focus at 6 cm. Color bar units are in decibels with respect to the maximum.

12-element Array: Individual Element Response

Individual element positions were examined to understand how the beam simulates shallow focusing. From this analysis, the beam was able to deliver energy up to a depth of 1 cm, remaining reasonably laterally localized at that depth. Based on these results, it can be concluded that the array elements can effectively sample the near field. In Figure 6, the X-Z plane shows that after the simulation reaches -6-decibels at 1 cm, the 6-decibel falloff occurs beyond 1 cm. This suggests that the beam maintains a strong focus up to 1 cm, which after, the signal weakens by 6 dB. relative to the strongest intensity. A depth of 1 cm is required because the width of the acoustic window is, on average, 1 cm.

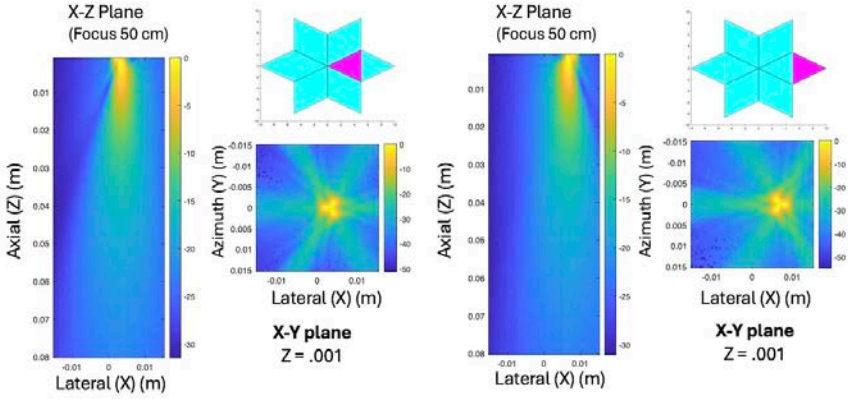


Figure 6.
Apodization of elements to characterize individual element response

12-element Array: Steering 10° Off Axis

The beam is steered 10 degrees off-axis in **Figure 7**. Grating lobes, or energy peaks off-axis to the main (steered) beam in **Figure 7**, are less than -6 decibels, indicating that the receive beam can be sufficiently steered and focused at different off-axis angles.[6] At Z=35 mm, the beam is effectively steered off axis.

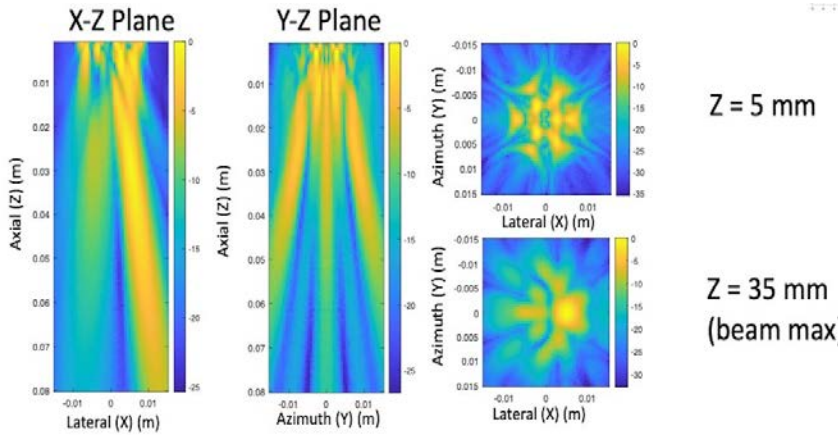


Figure 7.
X-Z, Y-Z, and X-Y plane simulations of the beam steered 10 degrees off axis. Color bar units are in decibels with respect to the maximum.

These results demonstrate the potential capabilities of a unique transducer array geometry designed for rapid stroke detection. A fundamental limitation was observed in the focal range, which is constrained to 35 mm. However, the array exhibited reasonable and useful beam shapes for the whole array whether on-axis, steered, or with individual elements. The array's basic capabilities include a focal depth of 35 mm, the ability to steer the beam off axis, and the capability to sample the near field. While the data displayed is preliminary, the final array design may require more elements to optimize its performance.

Future Work

In future work, the goal is to simulate a mock field of vessels at on-axis and off-axis locations to adjust the receive beam to detect a greater presence of flow at different angles. In this study we have only simulated the pressure fields. Vector velocity processing will also need to be conducted to improve sensitivity. An acoustic window-finding algorithm will need to be derived from experimental data. It should also be noted that the initial array geometries investigated here are not necessarily a final design; it may be necessary to increase the number of array elements to improve steering and focusing, or to more accurately measure flow density. This depends on the to-be-experimentally-derived algorithm, and how finely spaced it needs the array elements to be.

Conclusion

In summary, a simplified 2D array transducer can be constructed to provide TCD users with feedback on trends in underlying skull thickness and beam alignment with a major cerebral vessel to aid in optimal signal acquisition. This means that the 2D array can steer and focus the ultrasound beam in different directions and at varying focal depths. The Field II simulations of the 12-element triangular transducer displayed the ability to sample the near field (insonify shallow skull at different lateral locations) and control the beam focus angle (flow). However, the 6-element transducer design was less effective due to an insufficient number of elements along the axis through the origin. With the 12-element transducer design the beam could be focused and steered to different ranges ([1:1:6] mm) and angles (10°), making it a promising design that has the potential to improve LVO stroke detection at the point-of-care.

Acknowledgment

The authors sincerely thank Dr. Mark Rubin, Dr. Brent Hoffmeister, and Dr. Gianmarco Pinton for their collaboration on this project. The authors also acknowledge Kashta Dozier-Muhammad, Taylor Christian, Estefania Viloria, for their related work on this project. Thank you to Dr. Samuel Johnson, my father and primary inspiration behind my interest in medical research. The funding source for this project is the NIH/NIBIB grant R15EB034185.

References

[1] Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Cheng, S., Delling, F. N., Elkind, M. S. V., Evenson, K. R., Ferguson, J. F., Gupta, D. K., Khan, S. S., Kissela, B. M., Knutson, K. L., Lee, C. D., Lewis, T. T., Liu, J., … American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee (2021). Heart disease and stroke statistics-2021 update: A report from the American heart association. *Circulation*, *143*(8), e254–e743. https://doi.org/10.1161/CIR.0000000000000950

[2] Malhotra, K., Gornbein, J., & Saver, J. L. (2017). Ischemic strokes due to large-vessel occlusions contribute disproportionately to stroke-related dependence and death: A review. *Frontiers in neurology*, *8*, 651. https://doi.org/10.3389/fneur.2017.00651

[3] Sekerdag, E., Solaroglu, I., & Gursoy-Ozdemir, Y. (2018). Cell death mechanisms in stroke and novel molecular and cellular treatment options. *Current Neuropharmacology*, *16*(9), 1396–1415.

[4] Luijten, S. P. R., Wolff, L., Duvekot, M. H. C., van Doormaal, P.-J., Moudrous, W., Kerkhoff, H., Lycklama a Nijeholt, G. J., Bokkers, R. P. H., Yo, L. S. F., Hofmeijer, J., van Zwam, W. H., van Es, A. C. G. M., Dippel, D. W. J., Roozenbeek, B., & van der Lugt, A. (2022). Diagnostic performance of an algorithm for automated large vessel occlusion detection on CT angiography. *Journal of NeuroInterventional Surgery*, *14*(8), 794–800. https://doi.org/10.1136/neurintsurg-2021-017842

[5] Purkayastha, S., & Sorond, F. (2012). Transcranial Doppler ultrasound: technique and application. *Seminars in Neurology*, *32*(4), 411–420. https://doi.org/10.1055/s-0032-1331812.

[6] Barthez, P. Y., Léveillé, R., & Scrivani, P. V. (1997). Side lobes and grating lobes artifacts in ultrasound imaging. *Veterinary Radiology & Ultrasound: The Official Journal of the American College of Veterinary Radiology and the International Veterinary Radiology Association*, *38*(5), 387–393. https://doi.org/10.1111/j.1740-8261.1997.tb02104.x

Gerrell Dabbs is a Computer Engineering and Mathematics major with a minor in Computer Science at the University of Memphis. He is expected to graduate in spring 2026 and then attend graduate school in the following fall. In summer 2024, he was awarded the Honors Summer Research Fellowship and conducted research under Dr. Chrysanthe Preza. Through the fellowship, he presented his research at the National Conference for Undergraduate Research in spring 2025. Gerrell desires to perform further research in image processing and plans to continue researching through the Vertical Integrated Projects program.

Chrysanthe Preza is the Kanuri Professor and Chair of the Department of Electrical and Computer Engineering at the University of Memphis, where she joined in 2006. She received her D.Sc. degree in Electrical Engineering from Washington University in St. Louis in 1998. She leads the research in the Computational Imaging Research Laboratory. Since 2022, she is the founder and Director of the Vertically Integrated Projects (VIP) Program at the University of Memphis. Her research interests are imaging science, estimation theory, computational optical sensing and imaging applied to multidimensional multimodal light microscopy and hyperspectral imaging, and computational imaging enabled by deep learning. She received a CAREER award by the National Science Foundation in 2009, the Herff Outstanding Faculty Research Award in 2010 and 2015, and she was the recipient of the Ralph Faudree Professorship at the University of Memphis 2015-2018. Since 2018, she has been the recipient of the Ravi and Eleanor Kanuri Professorship. She was named Fellow of the SPIE in 2019 and Fellow of the Optica (OSA) in 2020. She served as Associate Editor for *IEEE Transactions on Computational Imaging*, as Topical Editor for Optica's *Applied Optics*, and as an Executive Editor for *Biological Imaging, Cambridge University Press.* She currently serves as an Editorial Board Member for *Journal of Imaging, Computational Imaging and Computational Photography section*

Gerrell Dabbs & Chrysanthe Preza

Noise Analysis for Hyperspectral Unmixing

Faculty Sponsor
Dr. Chrysanthe Preza

Abstract

Hyperspectral Imaging is a three-dimensional imaging technique which collects information from the electromagnetic spectrum. Hyperspectral images contain spectral signatures allowing for material identification and classification, making it useful for many fields. Hyperspectral imaging sees use in anomaly detection where there are several pipelines to utilize the imaging technique. An issue with these pipelines is their lack of generality, but a former graduate student of the CIRL successfully created a more general pipeline. Like others, their pipeline was created specifically for anomaly detection. However, there is interest in how it performs in purely noisy hyperspectral unmixing. In our investigation of the pipeline's behavior in this task, we find that the Sobel Operator would create solid color abundance maps which would cause a loss of information. As such, the operator was replaced with sharpening with varying degrees of success. However, there is still a need for a general solution.

Introduction

Images are indispensable in the sciences. They are a powerful tool when it comes to both representing and storing valuable information for different analytic tasks. The ability to store image data is also powerful as it allows for the analysis of scenes at both a different time and place than the original event recorded. By storing this information, we also can make use of computational techniques to extract features that would otherwise go unnoticed by the human eye. This field is known as image processing which is the act of performing different operations on images to analyze features in them. This includes the detection and enhancement of edges in an image, adjusting the contrast, and many other techniques. This can be seen in devices such as smartphones where it is responsible for our ability to search for information based on solely images or augmented reality, the act of superimposing objects onto reality using the camera on those devices.[6] Imaging processing sees further use in medical applications such as x-rays, or in surveillance, where people can be alerted of possible dangers. [4] Artificial intelligence makes use of imaging processing practices to effectively allow a computer to see the same way a human does in a field known as 'computer vision.' The proper utilization of image processing helps to streamline and improve daily life providing a permanent reason for the growth of the field.

The development of image processing requires understanding of the fact that not all images are the same in terms of data types. There are many applications of imaging processing that make use of standard RGB imaging which collects photos on the wavelengths of visible light and are the most typical data type when talking about an image. However, there exist other imaging methods which collect information over different wavelengths, including those which are not visible to the human eye. There are benefits to working with these other imaging types with a primary one being that they are more informative in some cases. Of these imaging types, we turn our attention to hyperspectral imaging, whihc is known for its ability to create information-dense datasets while allowing for the analysis and utilization of the special spectral properties for particular applications. These spectral properties allow for unique approaches to identifying what a scene contains while also doing so with a high degree of certainty than an imaging type such as standard RGB imaging cannot achieve. However, this does come at the cost of computational resources, with hyperspectral datasets being harder to handle. Despite this, it is worth working through the computational complexities and cost to utilize this imaging type.

Hyperspectral Imaging

Hyperspectral Imaging is an imaging technique which collects information from the electromagnetic spectrum rather than the visible light spectrum. This technique creates three-dimensional data by taking photos on the 'x' and 'y' axes at several different wavelengths denoted by 'λ'. The collection of photos is then 'stacked' on top of one another to create a 'data cube', a 3-dimensional representation of the data. In a hyperspectral image, pixel values represent reflectance which is the amount of light captured by the hyperspectral sensor at a given wavelength. While reflectance values are like color values or intensity values in visible light and grayscale images respectively, the analysis of reflectance values yields more information for identification when compared to the pixel value types.

Each unique material, or endmember, has a unique reflectance value across wavelengths known as a spectral signature. As this spectral signature is unique, there exists a particular spectral signature for each material that can be used to identify them, therefore making it possible to identify objects in a scene given there exists an acceptable ground truth or reference point for a material's spectral signature. For this reason, hyperspectral imaging is more practical than visible light imaging in applications focused on identification and classification.

Notably, there exists a similar imaging technique known as multispectral imaging. Multispectral imaging makes use of a similar concept to hyperspectral imaging in the sense that it is imaging over additional wavelengths as compared to just the visible light spectrum. However, hyperspectral imaging is still over more wavelengths than multispectral imaging. While multispectral imaging suffers from its lower wavelength count, it additionally suffers a lower spectral resolution when compared to hyperspectral imaging.[3] For the previously stated reasons, hyperspectral imaging sees usage in many fields such as astronomy, chemistry, agriculture, medicine, and surveillance. An example of a medical application is its use for tissue analysis, where it can be used to analyze certain properties of tissue. Another medical application is the analysis of absorption spectra to find the concentration and oxygen saturation of hemoglobin allowing for the identification of cancer hallmarks. [2]

Hyperspectral Unmixing

While hyperspectral imaging is a powerful tool on its own, it does suffer from a particular limitation. To ensure the accuracy of reflectance values, it is important that each pixel contains only one endmember, making that pixel pure. Otherwise, a mixed pixel will have a mix of reflectance values

within it, making it difficult to determine the material of the pixel. In this case, there is a need to 'unmix' the pixel to retrieve the reflectance of each individual material. The process of recovering the reflectance values is known as Hyperspectral Unmixing.

There are two approaches to unmixing: Linear and Nonlinear. Nonlinear is a more complex process, so linear unmixing will be used for an explanation. In the case of linear unmixing, the goal is to recover a linear combination of the endmembers which means that the coefficient of each endmember must be found. These coefficients represent the weighing of each endmember in a given pixel. If a pixel is pure, then the linear combination should be only one endmember with a coefficient of 1. If a pixel is mixed, then it should contain a noteworthy term for each endmember in the pixel with the coefficient for endmembers not contained in the pixel being 0 and therefore not affecting the pixel.

To perform hyperspectral unmixing, several algorithms are employed. Some of these hyperspectral unmixing algorithms are pixel purity index (PPI), fast iterative pixel purity index (FIPPI), and N-finder (N-FINDR). These algorithms exist to perform endmember extraction. It is important to understand that each algorithm performs differently in different cases so there is not a single best algorithm for general use. When unmixing, it is useful to have a visualization of the data. As such, there is a photo generated when unmixing known as an abundance map. An abundance map is a color scored image where the color at a given pixel indicates how much of a given endmember is contained at that pixel. This can be seen in Figure 1 where the warmer the color at a given pixel, the more that the material the abundance map is tracking is present in that pixel. For example, the third abundance map is tracking where there is water in the scene. The body of water is warm colored in the abundance map, showing that those pixels contain water. In the other abundance maps, the area where there was water in the third abundance map has a cool color, showing there is little to no water in those pixels.

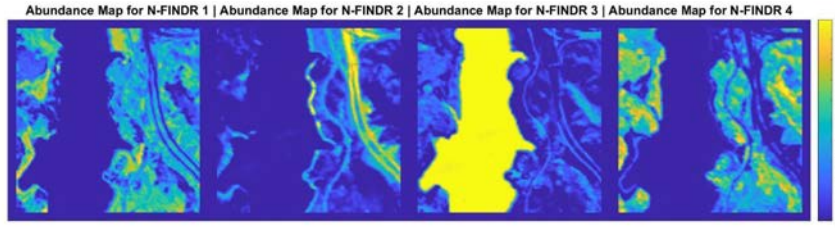


Figure 1.
Abundance Map Visualization

Focus of Study

This study is based on anomaly detection and a pipeline created for its enhancement when using hyperspectral processing techniques. Anomaly Detection refers to the process of finding outliers in a dataset and in the case of imaging, this would be to find out of place items. As described earlier, hyperspectral imaging allows for classification and identification of objects in a scene without the need for prior knowledge other than spectral signatures. As such, it is a powerful tool for the task and any advancements would help greatly in the field of surveillance.

The pipeline stems from a study conducted by a former graduate student in the CIRL [7]. The goal of his study was to create a pipeline that would improve hyperspectral anomaly detection along with outperforming pre-existing methods. The reason for this is due to the lack of generality in existing approaches to anomaly detections. Before his own contributions, pipelines either made a gaussian assumption, which assumed that anomalies were small and infrequent, or were computationally expensive with memory being noted as the main resource.[7] The student's method was an ensemble allowing for individual parts to be modified to produce better results. The parts of the method were band selection, gaussian blurring, and Sobel edge enhancements. This pipeline was successful in producing better results for anomaly detection, but it highlighted additional aspects to analyze such as the effect of different paint types on endmember extraction and how wrapped objects interact with the pipeline.

While there were different ideas to explore, there was one area of interest: noisy unmixing. The graduate student's pipeline had the goal of enhancing hyperspectral anomaly detection. The pipeline was not made with solely hyperspectral unmixing nor unmixing in noisy settings, but it makes mention of the importance of reducing noise. Noise naturally exists in hyperspectral images due to the resolution, but most popular datasets

have a reasonable level of it making it acceptable for hyperspectral tasks and testing. Simulating additional noise in these popular datasets would allow for an analysis of the pipeline under noise in a controlled and systematic way. As such, we had the goal of investigating if and how the pipeline handles noise and finding improvements in the case that it could not.

Testing the Original Pipeline

For our study, we used two known hyperspectral datasets. The first dataset is a popular dataset known as Jasper Ridge. The version of the dataset we are using is a 100 x 100 subset recorded at 198 wavelengths making the dimensions 100 x 100 x 198. It is known that there are 4 endmembers: Tree, Soil, Water, and Road. A representation of the dataset can be seen in Figure 2.

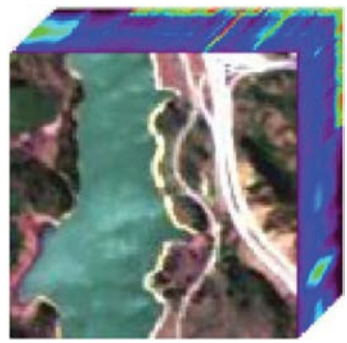


Figure 2.
Jasper Ridge Data Cube Visualization [5]

The second dataset is known as Cuprite, a picture of multiple minerals. The image is a 250 x 190 region recorded at 224 wavelengths making the dimensions 250 x 190 x 224. There are 14 endmembers, each a different mineral. However, an endmember count of 12 is accepted due to similarities between minerals. A representation of the dataset can be seen in Figure 3.

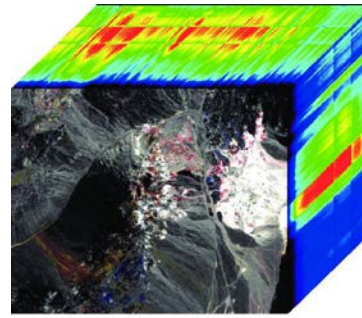


Figure 3.
Cuprite Data Cube Visualization [1]

Another factor to note is that some datasets have ground truths. A ground truth is simply an additional piece of data which indicates the absolutes of the dataset. It will inform the user what material each endmember is, the reflectance values at each wavelength, and the correct abundance map for each endmember. Furthermore, ground truths are split into a file for the abundance maps and a file for the reflectance values or endmembers. Jasper Ridge has an abundance map for both the reflectance values and abundance maps, but the file format for the abundance maps is not possible to work with for easy comparison. Cuprite has a ground truth for reflectance values, but not for its abundance maps. This makes it impossible to compare any experimental data related to abundance maps to absolute information. Additionally, any oddities with abundance maps from this dataset would not be apparent due to this missing data. The lack of ground truths is not a rare occurrence in Hyperspectral Imaging as it is not easy to produce ground truth data.

As noise is an integral part of our study, it must be quantified in some manner. Our metric for noise is known as Signal-to-noise ratio or, as it will be referred to, SNR. The higher the SNR, the lower the noise. It is worth noting that there exists some amount of noise naturally in experimental data, especially due to the resolution of hyperspectral sensors. As such, even when only a little noise is simulated, there will still exist noise from the original dataset. Nevertheless, we will treat the original dataset before noise simulation as the ‘noiseless’ dataset. The simulation and addition of noise was conducted through a user-defined MATLAB function which uses the pixel values in the image to generate an array of random numbers scaled to an input value with the same dimensions as the image to add to the image to create a noisy image.

Additionally, as we want to determine the impact of noise, we must use additional metrics. There are two metrics of value for this kind of study, Mean-Squared Error, which will be referred to as MSE, and Structural Similarity Index Measure, which will be referred to as SSIM. In imaging, MSE calculates the error between two images while SSIM calculates the similarity between same position pixels in two images. When it comes to evaluation of performance, the goal is to minimize the value of MSE while maximizing the value of SSIM. In MATLAB, there are built-in commands for the acquisition of these metrics. This is to say that MSE and SSIM values will be calculated between the noiseless abundance map for each endmember and their noisy counterparts at the different SNR values. This will allow us to see how similar the noisy processed results are to the noiseless ones.

To establish a starting point, the two datasets were passed through the pipeline so that the metrics could be computed. In all cases, the MSE and SSIM metrics were computed between the noiseless and noisy abundance maps of the given dataset.

Jasper	MSE					
SNR (dB)	22.35	24.57	26.61	29.62	35.64	38.65
1: Soil	4.270E-02	3.560E-02	2.940E-02	2.780E-02	1.880E-02	1.820E-02
2: Road	2.290E-02	2.240E-02	2.370E-02	1.890E-02	2.000E-02	2.230E-02
3: Water	5.000E-02	4.370E-02	3.220E-02	3.180E-02	2.390E-02	2.590E-02
4: Tree	2.540E-02	2.080E-02	1.670E-02	1.590E-02	1.390E-02	1.370E-02

Table 1.
Jasper Ridge (Original Pipeline),
MSEs for Individual Endmember #1-#4

In the case of the Jasper Ridge data, a few things can be noted (Table 1). On average, the endmember with the lowest MSE was the tree endmember, followed by the road, soil, and then water. The two highest MSE endmembers also see a large jump in their MSE values as the SNR decreases.

Jasper	SSIM					
SNR (dB)	22.35	24.57	26.61	29.62	35.64	38.65
1: Soil	8.70%	10.95%	13.06%	16.33%	22.55%	35.30%
2: Road	40.25%	43.21%	47.67%	48.79%	53.68%	48.22%
3: Water	36.72%	36.51%	40.85%	37.45%	43.46%	43.99%
4: Tree	18.67%	23.06%	38.02%	27.32%	36.34%	51.60%

Table 2.
Jasper Ridge (Original Pipeline),
SSIMS for Individual Endmembers #1-#4

Viewing **Table 2**, it can be seen that SSIM differs from MSE with the highest SSIM obtained for the road. Additionally, the tree and soil have the greatest change in their SSIM values at the highest SNR to the lowest SNR. With some exceptions, the increase of SNR results in the decrease of MSE and the increase of SSIM with the opposite also being true.

Cuprite	MSE					
SNR (dB)	27.44	30.45	33.46	36.47	39.48	42.49
1	9.917E-03	1.521E-02	8.152E-03	9.360E-03	7.187E-03	3.904E-02
2	1.894E-02	2.114E-02	1.904E-02	3.157E-02	2.549E-02	2.929E-02
3	1.459E-02	2.403E-02	1.529E-02	1.317E-02	1.818E-02	3.279E-02
4	7.757E-03	8.742E-03	1.697E-02	1.232E-02	1.049E-02	9.807E-03
5	2.922E-02	2.547E-02	2.603E-02	2.921E-02	2.676E-02	3.193E-02
6	2.017E-02	2.308E-02	2.275E-02	2.445E-02	2.002E-02	3.101E-02
7	4.922E-03	4.251E-03	5.705E-03	1.813E-02	3.861E-03	2.817E-02
8	2.135E-02	2.237E-02	2.541E-02	2.706E-02	3.045E-02	3.784E-02
9	2.128E-02	1.980E-02	2.495E-02	3.993E-02	2.191E-02	3.549E-02
10	5.497E-02	5.212E-02	5.032E-02	5.738E-02	5.319E-02	8.153E-02
11	4.583E-02	4.739E-02	4.548E-02	5.003E-02	4.975E-02	8.037E-02
12	4.151E-02	4.059E-02	6.239E-02	5.488E-02	4.407E-02	6.582E-02

Table 3.
Cuprite (Original Pipeline),
MSE for Individual Endmembers #1-#12

Cuprite	SSIM					
SNR (dB)	27.44	30.45	33.46	36.47	39.48	42.49
1	10.47%	9.14%	36.07%	39.54%	43.07%	23.36%
2	7.79%	6.78%	14.41%	9.11%	6.14%	7.60%
3	7.24%	6.26%	22.07%	25.88%	20.09%	6.67%
4	23.79%	30.38%	7.68%	26.15%	24.19%	23.97%
5	5.09%	17.42%	15.88%	11.04%	22.84%	3.94%
6	15.31%	11.99%	11.87%	9.76%	14.78%	7.75%
7	34.70%	40.92%	33.98%	30.85%	41.95%	13.28%
8	11.97%	19.23%	15.84%	11.32%	17.44%	5.42%
9	8.61%	20.16%	18.42%	2.90%	20.54%	10.73%
10	10.50%	5.97%	15.30%	18.72%	20.16%	8.38%
11	5.74%	6.13%	6.10%	5.25%	6.47%	1.70%
12	4.07%	4.24%	1.02%	1.87%	3.48%	1.21%

Table 4.
Cuprite (Original Pipeline),
SSIMS for Individual Endmembers #1-#12

When using the pipeline on Cuprite (Tables 3 and 4), endmembers are not labeled by their material as there does not exist a ground truth for the abundance maps. Cuprite also has a sizable number of endmembers, so lacking the ground truth makes it hard to extract data by inspecting the values of MSE and SSIM. Due to the large endmember count, it is helpful to take the MSE values for each endmember at a given SNR value and sum them together to create a sum of MSE Values for each SNR value resulting in a sum of errors that will be referred to as 'Sum MSE'. For SSIM, we can take an average of the SSIM values for each endmember at a given SNR value to create an average similarity value that shall be referred to as 'Average SSIM'. By doing this, it is easier to observe the effect of noise on the metrics without having to look through each individual endmember.

SNR (dB)	Sum MSE	Average SSIM
27.44	2.905E-01	12.11%
30.45	3.042E-01	14.89%
33.46	3.225E-01	16.55%
36.47	3.675E-01	16.03%
39.48	3.114E-01	20.10%
42.49	5.031E-01	9.50%

Table 5.
Cuprite (Original Pipeline),
Summed MSE and Averaged SSIM

Condensing the data into sums and averages, as seen in Table 5, makes it clear that the behavior of the pipeline on Cuprite is similar to its behavior on Jasper Ridge with the main difference being that the highest SNR's results are not better than the results at the lower SNRs. We once again see that an increase in SNR results in a decrease in MSE and an increase in SSIM.

When running the pipeline on these datasets, it was determined that Sobel Operation did not work as its inclusion would result in the generation of solid color abundance maps, meaning that the maps were indicating that the scene consisted of entirely one endmember, and therefore rendering the abundance maps useless. This was confirmed as when removing it, abundance maps returned to normal. As such, the Sobel Operator was removed. However, the removal of the Sobel operation leaves out a valuable part of the pipeline, calling for a need to replace it. To try to remedy this, sharpening was implemented in its place by creating a sharpening mask or matrix to then apply to the image. Sharpening performs a similar action to Sobel-Edge Enhancements while it avoids producing solid color abundance maps.

Testing the Modified Pipeline

The testing for the modified pipeline is the same as for the original pipeline.

Jasper	MSE					
SNR (dB)	22.35	24.57	26.61	29.62	35.64	38.65
1: Soil	4.990E-02	4.010E-02	3.060E-02	2.780E-02	1.320E-02	1.370E-02
2: Road	2.540E-02	2.160E-02	2.070E-02	1.530E-02	1.360E-02	1.520E-02
3: Water	4.940E-02	3.970E-02	2.930E-02	2.610E-02	1.830E-02	2.040E-02
4: Tree	3.250E-02	2.040E-02	1.450E-02	1.210E-02	8.300E-03	8.600E-03

Table 6.
Jasper Ridge (Modified Pipeline),
Individual Endmember MSEs

The modification to the pipeline causes Jasper Ridge's endmembers to have lower MSE numbers at higher SNR values (Table 6). For 38.65 dB and 35.64 dB, each endmember sees a minimum of a 21% decrease in MSE with some seeing 30% or greater. At 29.62 dB, the soil does not benefit while the others see a decrease of at least 18%. However, at lower SNR values, there is an increase in MSE in the modified pipeline compared to the original pipeline. For 24.57 dB, the soil sees a 13% increase in MSE while the others still see slight decreases. At 22.35 dB, only water decreases by 1% while the others increase by at least 11%.

Jasper	SSIM					
SNR (dB)	22.35	24.57	26.61	29.62	35.64	38.65
1: Soil	17.12%	26.16%	32.73%	37.77%	54.44%	64.83%
2: Road	32.28%	46.31%	49.10%	57.44%	66.59%	72.49%
3: Water	27.80%	42.27%	38.42%	41.57%	55.01%	59.46%
4: Tree	34.16%	42.27%	47.41%	50.78%	60.47%	76.06%

Table 7.
Results from Jasper Ridge (Modified Pipeline),
Individual Endmember SSIM

For SSIM (Table 7), the results are similar to MSE. At 38.65 dB and 35.64 dB, there is a minimum increase of 17% and 1%. All other SNR values see a decrease in either one or two endmembers. From 29.62 dB, soil always performs worse, starting with a 23% decrease in SSIM and reaching a 57% decrease at 22.35 dB. The road endmember also does worse, but only at 22.35 dB. The water and tree endmembers always see an increase in SSIM.

Cuprite	MSE					
SNR (dB)	27.44	30.45	33.46	36.47	39.48	42.49
1	5.039E-02	4.965E-02	2.730E-02	1.866E-02	1.266E-02	1.134E-02
2	1.910E-02	3.858E-02	3.784E-02	7.321E-02	4.941E-02	7.420E-02
3	1.801E-02	4.093E-02	5.173E-02	3.505E-02	7.720E-02	4.067E-02
4	1.740E-02	7.959E-03	2.269E-02	1.650E-02	1.002E-02	5.220E-02
5	3.090E-02	3.484E-02	7.552E-02	2.906E-02	1.097E-02	3.754E-02
6	2.022E-02	1.505E-02	1.822E-02	4.091E-02	3.890E-02	2.178E-02
7	3.102E-02	1.660E-02	3.528E-02	2.984E-02	8.615E-03	5.177E-03
8	3.391E-02	2.014E-02	1.345E-02	3.621E-02	1.908E-02	5.154E-02
9	2.869E-02	2.014E-02	3.027E-02	7.916E-02	3.492E-02	3.856E-02
10	4.596E-02	4.338E-02	3.252E-02	7.941E-02	9.119E-02	5.043E-02
11	4.169E-02	3.864E-02	1.130E-01	1.162E-01	5.172E-02	2.750E-02
12	2.409E-02	8.129E-02	3.046E-02	3.523E-02	3.004E-02	3.468E-02

Table 8.
Results from Cuprite (Modified Pipeline),
MSEs for Individual Endmember #1-#12

Cuprite	SSIM					
SNR (dB)	27.44	30.45	33.46	36.47	39.48	42.49
1	0.99%	1.56%	0.10%	3.09%	15.68%	6.18%
2	30.90%	13.97%	14.60%	-17.75%	-13.33%	12.54%
3	29.10%	23.81%	23.98%	36.90%	-8.23%	34.13%
4	15.72%	43.11%	4.16%	8.61%	29.07%	1.40%
5	19.10%	5.32%	-10.72%	25.26%	59.54%	12.25%
6	27.66%	52.45%	50.51%	-6.38%	0.16%	40.53%
7	-2.78%	5.44%	-0.54%	15.18%	21.91%	38.00%
8	-0.29%	40.70%	41.20%	3.84%	51.77%	-8.58%
9	39.92%	47.23%	12.57%	-9.27%	43.25%	33.16%
10	35.36%	31.61%	45.12%	16.10%	9.81%	39.25%
11	45.52%	43.30%	-4.61%	-5.51%	50.69%	53.10%
12	64.29%	-7.24%	47.29%	63.03%	58.07%	47.57%

Table 9.
Results from Cuprite (Modified Pipeline),
SSIMs for Individual Endmember #1-#12

For Cuprite (Tables 8 and 9), there is ample variance in the behavior at different SNRs. We use the same idea from earlier and condense our data into sums of errors and averages of similarities.

SNR (dB)	Sum MSE	Average SSIM
27.44	3.614E-01	25.46%
30.45	4.072E-01	25.11%
33.46	4.883E-01	18.64%
36.47	5.895E-01	11.09%
39.48	4.347E-01	26.53%
42.49	4.456E-01	25.79%

Table 10.
Results from Cuprite (Modified Pipeline),
Summed MSE and Average SSIM

In Cuprite (Table 10), the MSE increases by at least 24% for all SNR besides 42.49 dB. At that SNR, there is an 11% decrease in MSE. SSIM

also increases at all SNRs except for 36.47 dB where it sees a 31% decrease. For all others, it sees at least a 13% increase and at most a 171% increase at 42.49 dB.

Discussion

As the data displays, there are improvements in SSIM and MSE at higher SNRs while lower SNRs see worsening results. This goes back to the graduate student's point on generalization and how it is a common problem among anomaly detection pipelines. There are two focus points when looking at the data and considering further questions: The Sobel Operator and the endmembers that only saw benefits from sharpening.

The Sobel Operator works by convolving the image with different matrices to determine where there is a sharp change in pixel values and therefore an edge. By adding simulated noise in random places in the scene, sharp changes exist at more points than just object edges. This results in the addition of extra edges in the scene. This is likely why solid color abundance maps were produced whenever the Sobel Operator was used on the scene. Another issue with removing the Sobel Operator is that it removes an additional layer of smoothing done by the operator itself.

The endmembers that only saw an improvement from sharpening are the water endmember for both MSE and SSIM, and the tree endmember only for SSIM. These are both endmembers which have many connected parts throughout the scene with water being one large body, and the tree being four sizable parts opposed to the scattered nature of the soil or the small road. It is notable that the road endmember in Jasper Ridge performed well at all SNRs except for the lowest, and that the road is comprised of roughly 3 pieces. This may indicate that the main issue in noisy unmixing arises from the inability to clearly define edges in the same way the Sobel Operator is likely to enhance the noise as stated prior.

The lack of generalization A possible solution to the issue is to investigate general denoising for hyperspectral imaging. One common approach to the need for denoising is found in machine learning and artificial intelligence where different models can be constructed. These methods can be either model-based or learning-based with both having their own requirements to work with.[8] In further study, this may yield the best result.

Conclusion

Imaging a prevalent tool in many fields yields useful information. Hyperspectral Imaging is an imaging method capable of creating informa-

tion-dense datasets compared to other methods such as RGB or visible light imaging. Due to the nature of hyperspectral imaging, it is also prevalent in anomaly detection where many pipelines make assumptions about a scene to produce results. In prior work, a pipeline was created that was able to produce better results in generalized settings unlike pipelines of the time. While the pipeline succeeded in its immediate goal, there was interest in how it performed in purely unmixing particularly cases with noise. To test this, noise was simulated through MATLAB commands. Testing revealed that the Sobel Operator did not respond well to noise but could be removed to process images regardless. The removal of the operator created a need for a sufficient replacement and for that sharpening was tried. Sharpening was able to improve some cases, but did not generalize causing a return to the original issue stated in the former graduate student's study. The pipelines prior to their contribution made assumptions about the given data to enhance the task at hand. However, these assumptions hurt the generality of the pipelines, resulting in varying results across different datasets. As there is a need for a more general solution, an approach to solving this problem could be found in the domain of machine learning and artificial intelligence models that are able to adapt based on the given dataset.

Acknowledgement

I would like to thank the Helen Hardin Honors College for enabling me to do research through their Summer Fellowship. I also would like to thank Dr. Chrysanthe Preza for being my research advisor and allowing me this opportunity.

References

[1] Kaya, S. (2017). A comparative analysis of classification methods for hyperspectral images generated with conventional dimension reduction methods.

[2] Lu, G., & Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of biomedical optics, 19* (1), 10901. https://doi.org/10.1117/1.JBO.19.1.010901

[3] Niu, B., Feng, Q., Chen, B., Ou, C., Liu, Y., & Yang, J. (2022). HSI-TransUNet: A transformer based semantic segmentation model for crop mapping from UAV hyperspectral imagery. *Computers and Electronics in Agriculture, 201,* 107297. doi:10.1016/j.compag.2022.107297

[4] Sharma, P., & Singh, R. (2022). *Application of Image Processing in Real World.*

[5] Song, D., Sun, N., Xu, M., Wang, B., & Zhang, L. (2020). Fast Unmixing of Noisy Hyperspectral Images Based on Vertex Component Analysis and Singular Spectrum Analysis Algorithms. *Canadian Journal of Remote Sensing, 46*, 1–15. doi:10.1080/07038992.2020.1726735

[6] Thabet, R., Mahmoudi, R. & Bedoui, M.H. (2014). Image processing on mobile devices: An overview. doi:10.1109/IPAS.2014.7043267

[7] Younis, M. (2023) "Investigation of hyperspectral data unmixing and its impact on the task of anomaly detection." *Electronic Theses and Dissertations*. 3343. https://digitalcommons.memphis.edu/etd/3343

[8] Zhang, T., Fu, Y. & Zhang, J. (2022). Guided Hyperspectral Image Denoising with Realistic Data. *Int J Comput Vis, 130*, 2885–2901. https://doi.org/10.1007/s11263-022-01660-2

Hema Dalavayi graduated *summa cum laude* from the University of Memphis in fall 2024 with a major in Computer Science and a second major in Mathematics. She was a student marshal for the College of Arts and Sciences, awarded for the highest GPA (4.0/4.0) in her overall undergraduate studies. During her undergraduate studies, she was actively involved in several research projects and presented her findings at the 19th & 20th Annual Computer Science Research Symposia in spring 2024 and 2025. She also pursued an accelerated BS/MS program and has been admitted into the PhD program in Computer Science at the University of Memphis in spring 2025 under Dr. Max Garzon.

Max Garzon is professor of Computer Science and Bioinformatics at the University of Memphis. The main theme of his research is biomolecule-based computing and applications to areas such as nanotechnology, self-assembly, semantic indexing and data science. He has served as TPC and organizer of many scientific conferences and professional meetings in these areas. He has been a Fulbright scholar, invited visiting professor and guest scientist at a number of research institutions around the world and enjoys interdisciplinary work on problems of common interest.

Hema Dalavayi & Max Garzon

On the Randomness of Evolutionary Events along Lineages:
Two Case Studies

Faculty Sponsor
Dr. Max Garzon

Abstract

Evolution first came into biology through the seminal work by Charles Darwin in the 1800s. Evolutionary events have often been seen as random, with genetic changes caused by factors like transcription, radiation, and environmental influences. This paper aims to explore whether evolutionary events are entirely random by examining three prediction problems for fishes of the Family *Cichlidae* (C), and bacteria from the genus *Helicobacter* (H). Predictions concern the divergence time of an organism (when its species first appeared on Earth) in a biological lineage, the difference in divergence time between two organisms, and a proxy for the next organism in the lineage. Neural network models were trained using machine learning techniques that yield relative errors as low as 5%, 6%, and 57% in their predictions, respectively. These results suggest that while evolution is influenced by random factors, some specific events may be predictable, adding evidence to the argument that evolutionary events are not entirely random.

Introduction

As pointed out by Dobzhansky, "nothing makes sense in biology except in the light of evolution" (Dobzhansky, 1950). The theory of evolution is a cornerstone of modern biology. The idea of evolution first came into biology through the seminal work by Charles Darwin (Darwin, 1859). Darwin proposed that species evolve over time through the process of natural selection, where individuals with traits suitable for their environment have a better chance to survive and reproduce, by passing these traits on to future generations. This idea revolutionized biology by providing a framework that explains how organisms on Earth change over time despite (perhaps drastic) changes in their environment. Darwin's theory challenged the theological views that were prevalent during that time. First, many theologists believed that life on Earth was static and that they were unchanged and fixed as perfect creations by God (Paley, 1802). Paley argued that each species has a fixed purpose in life as designed by a divine Creator, with every part of an organism functioning to serve a specific role in the natural world. This view supported the belief that species were immutable and had remained unchanged since their creation. Unlike Paley's view of a static world, Darwin argued that species change over time in response to environmental pressures. This challenged the longstanding belief in the static world and incurred a lot of criticism from both the religious and the scientific communities. However, after much debate and controversy over the years, this theory of evolution was gradually accepted by the scientific community and has become prevalent today.

Evolutionary events have often been assumed to be entirely random by biologists because genetic changes (e.g., variations and mutations) depend on changes occurring in an organism's DNA due to many internal factors affecting individuals. Also, evolution is also influenced by environmental factors such as geography, climate and available resources. Now, statements of this sort are not very meaningful without a noncircular and rigorous definition of randomness. There are two important definitions of randomness, statistical and computational. According to Wikipedia [9], statistical randomness refers to a sequence of numbers that contains no recognizable patterns or regularities, making it impossible to predict the next number based on the previous elements in the sequence. Likewise, computational randomness refers to sequences where the elements cannot be predicted or generated by any algorithm or computer program. In other words, it is unpredictable to the extent that no computer, no matter how powerful, can predict the next element from the previous ones. Although important for our research question, the purpose of this paper is not to pro-

vide a precise definition of randomness, hence for the sake of this study, random will refer to outcomes that are unpredictable. Therefore, although Darwin never said so explicitly, most biologists have come to assume that evolutionary events are unpredictable, as it is accrued by random genetic mutations and interactions of an organism with its environment that accumulate over space and time.

On the other hand, recent studies have shown that genetic variations follow noticeable patterns though they are indeed influenced by factors such as the environment, natural selection, and evolution. For instance, (Wortel et al, 2023) states that two key factors influencing evolution are genetic elements, such as mutation bias and epistatic interactions, as well as ecological factors. Genetic factors, particularly in large populations, provide more variation for natural selection to act upon, while rapid environmental changes complicate predictions due to their complexity. This research laid the groundwork for exploring the predictability of evolution through genomic data. In addition, a case study by (Mas et al, 2020) made observations on how natural selection affected allele frequency changes in the plant of the species Mimulus guttatus. The study demonstrated that genetic models could effectively predict evolutionary outcomes when fitness measurements were available, suggesting that evolutionary changes are not entirely random but rather guided by specific genetic and ecological pressures. This fact aligns with the evolving view that, while genetic mutations are random, selection pressures can steer evolutionary processes in predictable directions in a population. These discoveries suggest that evolutionary events could be influenced by more than just random mutations pointing to the presence of underlying patterns in DNA that guide evolutionary changes.

The primary aim of this paper is to examine at a deeper and narrower level whether evolutionary outcomes are entirely random. While evidence in the previous paragraph suggests that some aspects of evolution can be predicted based on genetic variations on populations, this paper aims to explore how un/predictable evolutionary changes may be for individual organisms or species and furthermore, based only on their DNA. It will investigate, in particular, the extent to which DNA contains necessary information to make certain features of evolutionary events predictable, contrary to what is expected in a purely random process. The question is precisely addressed by solving 3 prediction problems for two (2) lineages: fishes of the family of Cichlidae (C) and bacteria from the genus HelicoBacter (H). Here, a lineage L is a sequence of individual organisms $(x_1, x_2, \ldots, x_N)$ that have evolved from a single common ancestor x_0. Such

prediction problems have not been considered because evolution in biology has been regarded as essentially unpredictable. Cichlids are eukaryotes known for their diversity and adaptability, making them very appropriate for studying evolutionary questions. Analogously, HelicoBacter is a genus of prokaryotes in the domain of bacteria, that has been studied for their role in human health and their potential for rapid evolutionary changes due to their short generation times. For this study, a prokaryote and a eukaryote lineage will be considered to formulate precise definitions of what is being predicted before the prediction is made, to avoid the circularities typical in the subject, as follows.

Defining a prediction problem PREDICTION (L, f) precisely for a lineage L requires a specification in advance of what is being predicted based on what information (Garzon et al, 2022). The evolutionary relationship in a lineage L such as Cichlidae fishes (C) and HelicoBacter (H) is modeled as a function f that maps organisms in a lineage L in a population to what is being predicted, usually a numerical value. The predictions concern the divergence time of a species, i.e. how long ago in the Earth's historical past did the species of an organism appear on Earth since the appearance of its immediate previous ancestor(s) in the lineage. The organisms in the taxon L will be represented by segments of their genomes (e.g. certain genes) and given in a feature vector (x_1, x_2 ... $x_{t)}$ of DNA sequences as proxies to represent each organism. The prediction problem is defined by a series of instances, each with the input necessary on which to base a prediction, and corresponding questions regarding what the prediction is about. A solution to the problem is some single device (e.g. a computer program) providing the right answers to the questions being asked for every single instance, regardless of how the correct answers may be changing from one instance to the next.

(a) [DIVTIME] PREDICTING DIVERGENCE TIME (L, f)

INSTANCE: a vector $\mathbf{x}_t = (x_1, x_2, \ldots, x_t)$ representing the previous species' DNA in the lineage up to time t.
QUESTION: What is the divergence time of organism $x_{t+1} = f(\mathbf{x}_t)$, i.e., when the next organism $x_{t+1} = f(x)$ in the lineage appears on Earth?

In biological phylogenetics, this time is usually measured in terms of millions of years ago (Mya) and is usually estimated using traditional methods like fossil records or carbon dating for a given organism (there is an inherent inaccuracy in these estimates given the nature of the methods used, e.g., radioactive decay in the fossil (Carleton, 2018).)

(b) [DDIV] PREDICTING DIFFERENCE IN DIVERGENCE TIME (L, f)

INSTANCE: Two organisms (x_{t-1}, x_t) in L
QUESTION: What is the difference in divergence times?

(c) [NEXTP] PREDICTING NEXT PROXY (L, f)

INSTANCE: a vector $\mathbf{x}_{[t-4,\ t-1]} = (x_{t-4}, x_{t-3}, x_{t-2}, x_{t-1})$ representing the 4 previous species' in the lineage L.
QUESTION: What is the next element xt= f($\mathbf{x}_{(t-4,\ t-1)}$) in the lineage at time t+1?

In this prediction problem, the goal is to predict the genomic signature of the next organism, not necessarily its exact DNA sequence. (With the predicted genomic signature, the exact original DNA sequence cannot be reconstructed).

In terms of difficulty, PREDICTING DIVERGENCE TIME (L, f) was the simplest problem of the three because given only the DNA of one organism, the solution needs to extract/ predict a number representing the divergence time. At first sight, this problem appears impossible to solve. Second, PREDICTING DIFFERENCE IN DIVERGENCE TIME (L, f) is a step deeper than the previous one because, in principle, a model does have to analyze the DNA of two different organisms to predict the difference in the divergence times, either individually or jointly. Finally, the most difficult problem was PREDICTING NEXT PROXY (L, f) because the model had to analyze the DNA in numerical form, for four (4) different organisms and predict the genomic signature of the next organism in the lineage, as it occurred in the Earth's history.

This paper is organized as follows. Section 2 provides a high-level characterization of data science and machine learning techniques to be used to provide solutions to these problems. Section 3 describes the solutions obtained for each problem and a summary and an assessment of the results. Finally, Section 4 presents a discussion of the findings from the research.

Methods

The traditional/conventional approach to addressing biological problems relies on qualitative observations or current or past life and their recordings to understand life, e.,g. as taxonomies of life as it exists now, or

hypotheses about the predecessors of that life in the reproductive cycle. However, solving prediction problems requires further analyses of the observations (data) to extract useful information (Garzon et al, 2022). This methodology has been refined by the emerging fields of data science and machine learning that will be used in this paper to handle these prediction problems. This methodology is structured into three distinct phases: observation of the phenomenon behind the problems and data collection, data preprocessing, and the development of machine learning models, such as neural networks, to obtain the solutions.

Data Science and Machine Learning

Data science is a field developed to address computational problems in ways that differ from traditional sciences, focusing on observing and gathering data about a phenomenon of interest rather than relying on deep, complex analyses to gain a deep understanding of a phenomenon, the conventional approach in science. With the advent of computer science and the information age, tools have been developed to analyze these huge amounts of data regardless of their specific domains (Garzon et al, 2022). This technique has enabled the identification of intricate patterns and trends across vast amounts of data which would have been a nearly impossible task to detect using the traditional methods. Thus, data science methods and tools have become a pivotal tool in solving many kinds of problems by leveraging machine learning techniques to identify patterns, make predictions, and optimize processes based on just data across a wide range of industries.

For each of the three problems defined above, it is essential to collect appropriate data to address them. In the data science approach, a computational tool (i.e. neural network) is then used to process the data, identify patterns, and generate accurate predictions based on the input. Out of the many types of tools available in machine learning to solve a problem, a neural network is particularly suited for tasks involving complex patterns in a dataset. A neural network (NN) is a model inspired by the human brain and consists of a set of (artificial) neurons, which are simple units that process local information. Each neuron (i) has a range of activation values Ai, along with a vector Wi of synaptic weights wi,j, where each weight represents the strength of the connection between two neurons i and j. In addition, each neuron is endowed with an activation function with a domain of real numbers and a codomain Ai of activation values that the neuron can assume and output after applying the activation function $\sigma_{i,}$ like the sigmoid function $\sigma(y) = \frac{1}{1+e^{x}}$. Typically, in a neural network,

each neuron computes the net input by taking the sum of the products of the activation values (a_j) and the synaptic weight $w_{i,j}$ of the neuron. This net input is then passed through the activation function σ_i which outputs a value within a specified range, such as [0, 1] or [-1, 1], depending on the function used. The resulting output becomes the activation value for the next time step. This process continues until the output layer is reached, where the final activation values produce the final output (e.g. a prediction associated with the inputs clamped to a pre-designated set of input neurons.)

Among the infinite number of possible neural networks, the correct network capable of solving a given problem is usually identified by a learning algorithm. In supervised learning, this algorithm enables the network to adjust its weights and activation functions by learning from labeled data, i.e., data containing the appropriate answers in advance. In this approach, the network is "supervised" by the labeled output (a.k.a target label) during the training phase. This means that for each of the input vectors in the training set, the correct output is already known so the goal of the network is to learn the relationship between the input and output by adjusting its weights so that with that knowledge it can make more accurate predictions on unseen data later. The neural network is deemed to make a correct prediction when its output matches the labeled data. If not, then the learning algorithm adjusts the parameters (weights) of the NN so it will a better chance to produce the correct output. Before starting the training phase, the learning algorithm splits entire dataset into a training and testing set, based on a parameter fixed by the researcher (typically 70%/30% or 80%/20% for training/testing), allowing the NN to be tested to evaluate its performance after the training is completed. The learning algorithm requires a structure of the neural network to be trained (the architecture) such as the number of layers, the type of activation function to be used and the data. After that, the learning algorithm starts with random weights to get a candidate network to improve during training. Then the learning algorithm processes each data point by feeding it as an input feature vector to the candidate network, which then does through its internal working as explained above and produces an output (this is known as the forward pass.) If the candidate network produces an output that does not match the labeled data, that means the model is not performing well, and the learning algorithm needs to adjust the weights.

Among the many types of learning algorithms available in supervised learning, the most efficient and popular is the backpropagation algorithm, which optimizes the weights of the neural networks based on the error

difference between the predicted and actual outputs. In the forward pass, if the output does match the labeled data, then the network proceeds to the next data point and continues the process. On the other hand, if the output does not match the labeled data, then the network takes on a backward pass where the error is propagated backward through the network. This backward pass works in a way such that the changes in the weights of the neurons are proportional to how much they have contributed to the incurred error. During this pass, the error is repeatedly propagated to one layer behind, penalizes the neurons and prompts change in their weights until it reaches the input layer. The learning algorithm repeats the whole process with other data points to complete one epoch. The user can set a stopping condition for the learning algorithm, such as specifying the number of iterations (epochs) to run or setting an accuracy threshold for accumulated errors, at which point the learning process stops.

Once this training phase is complete, the learning algorithm shifts to the testing phase where the trained NN is evaluated using the testing dataset that it has not seen before. The NN processes each input and makes a prediction. Then the prediction is matched by the learning algorithm against the target label and if it matches, then the prediction is marked as correct, else an error computed as the difference. Once the neural network has made all the predictions on the testing dataset, an accuracy measure is obtained. If the network's performance meets the desired criterion of quality, it is ready to effectively solve the defined problem; else the training is repeated with different parameters to obtain a better performing model.

Data Gathering and Pre-Processing

DNA encodes for critical information required to develop and sustain life in every living organism. Therefore, DNA sequences are the most appropriate data to obtain for solving the problem due to the deep structure of the DNA spaces. The lineages data was downloaded from GenBank (Sayers et al, 2022) as DNA sequences. Each data point was a DNA sequence of a specific gene. For the Cichlidae, the genes NADH (dehydrogenase) or COI (Cytochrome Oxidase I) of an organism in the lineage were used. NADH dehydrogenase is part of the mitochondrial genome which plays a crucial role in cellular respiration and is known for evolving at a moderate rate. The COI gene is also part of the mitochondrial genome and is known for evolving at a faster rate compared to other genes, making it useful for studying evolutionary events, such as the divergence times.

For the Cichlidae and the HelicoBacter lineages, the sequences were downloaded for their specific genes using the accession numbers (unique

identifiers) (Garcia and Colorado, 2024). For the Cichlidae-NADH, DNA sequences for 44 NADH genes and 36 COI genes were downloaded. Likewise, for the genus HelicoBacter(H), the DNA sequences of 69 organisms for genes NixA and trpC across 6 different species. The divergence times for each organism in both lineages have been provided by (Garcia and Colorado, 2024). **Tables 1** and **2** summarize these data sets.

Datapoint (NADH + COI)	Time Frame Mya: million years ago.	Accession Number NADH	Accession Number COI
Cyprichromis leptosoma	1.3	AY740381	AB915464
Eretmodus cyanostrictus	2.0	DQ055010	KU194153
...			
Tanganicodus irsacae	2.0	DQ055007	HQ533431
Altolamprologus calvus	0.95	EF462256	KU194199

Table 1.
Typical data points from various organisms
in the Cichlidae Fish (C) lineage.*

Datapoint (NixA+trpC)	Time Frame Mya	Accession Number
H. pylori	0.521	AWNG00000000
H. ailurogastricus	N/A	CDMH00000000
H. felis	0.689	FZKF00000000[b]
H. suis	0.2	FZKI00000000[b]
H. acinonychis	0.049	FZMD00000000[b]
H. salomonis	1.41	FZKZ00000000[b]
... followed by 64 more rows		

Table 2.
Typical data points from 6 species of organisms
in the HelicoBacter (H) lineage.*

*The accession numbers were used to obtain the DNA sequences for the organism from the GenBank [8]

To solve a problem in data science using neural networks, the DNA sequences cannot be directly given as inputs to a neural network because it only takes numerical feature vector values as inputs. Therefore, these DNA sequences were transformed into feature vectors by selecting a non-cross-hybridizing (nxh) basis and using the genomic signature of the DNA proxy for individuals in the lineage as a feature vector. Using the deep structure of DNA spaces (Garzon et al, 2022), three nxh bases were selected to transform a DNA sequence into numerical feature vectors. Given that there are multiple probes of length m in each nxh basis, each DNA sequence of an

organism was shredded into nonoverlapping fragments of length m. Now, each shred is compared with each probe to see whether they hybridize or not. Finally, for each probe, the count of number of shreds hybridized with the probe is normalized by dividing it with the total number of shreds. Finally, this results in a numerical vector for each probe that represents the relative frequency of hybridization between the DNA sequence and the probes in the nxh basis. This vector, called the genomic signature, captures essential information in the DNA sequence of an organism. (Garzon et al, 2022). For the Cichlidae-COI and NADH genes, 4mP3 and 5mP6 were used to convert the DNA sequences into numerical vectors. Here, 4mP3 refers to a probe of length 4 and P3 refers to the length of the vector. Table 3 summarizes the dimensionalities of the datasets.

Lineage - Genes	nxh basis	Feature vectors (length)	Size of the dataset
Cichlidae - NADH	4mP3	3 → 2 *	44
	5mP6	6	44
	4mP3+5mP6	8	44
Cichlidae - COI	4mP3	3→2 *	36
	5mP6	6	36
	4mP3+5mP6	8	36
HelicoBacter-NixA+trpC	4mP3	3→2 *	69
	5miC3Mg	3→2 *	69

Table 3.
Summary of feature vector lengths and
number of data points for lineages C and H.

* The original vectors have been reduced [3 → 2] for training purposes.

Training Neural Network Solution

With these tools in hand, it was possible to train a neural network (NN) to solve the target problems. During the backpropagation training phase, the feature vectors in Section 2.2 were given as inputs to the neural network (the number of input neurons to the NN was changed to match the length of the feature vectors.) The training phase halted once the predefined stopping conditions were met. After training the NN multiple times, the NN's weights converged after 10,000 epochs. The NN solutions are shown in Figures 1, 2, 3.

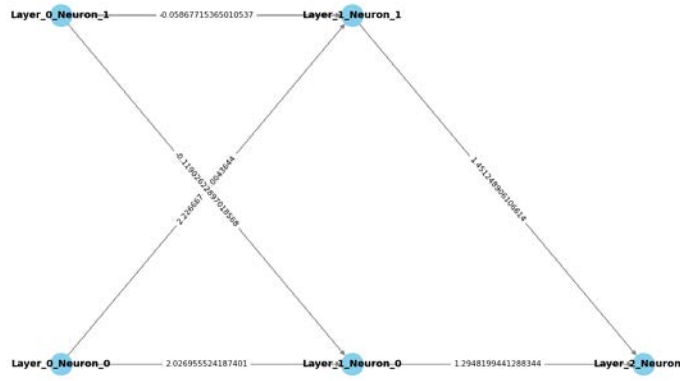


Figure 1.
Neural Network solution for
the problem of PREDICTING DIVERGENCE TIME (L, f).

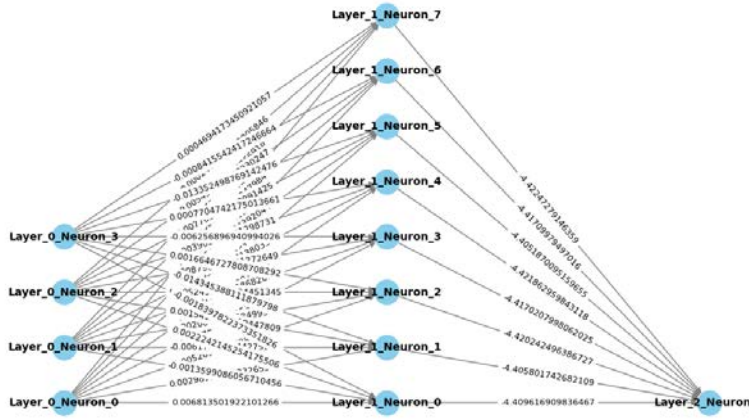


Figure 2.
Neural Network solution for the problem of
PREDICTING DIFFERENCE IN DIVERGENCE TIME (L, f).

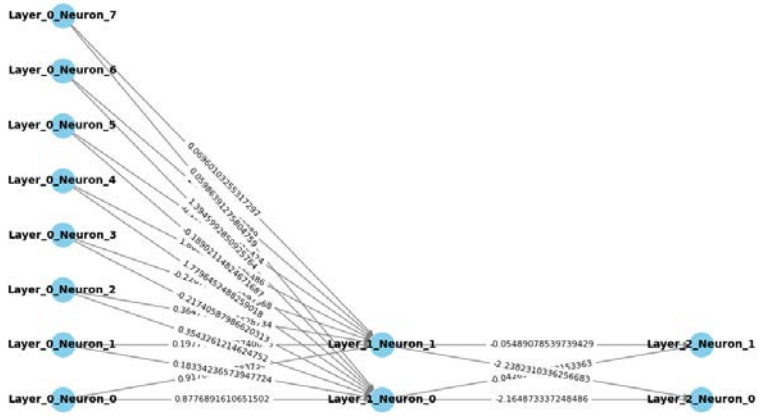


Figure 3.
Neural Network solution for the problem of
PREDICTING NEXT PROXY (L, f).

To assess the robustness and reliability of the NN model, validation techniques were applied to evaluate its performance and ensure its generalization ability on unseen data. Once the training phase was completed, the learning algorithm moved to the testing phase of the trained NN, where unseen data was given as input to generate predictions. Then the predictions were compared with the labels (target values) and an accuracy assessment was made. If the accuracy threshold set by the user was reached, the NN model was accepted as a solution to the prediction problem.

Results and Assessments

The solutions to the problems are trained NN models that predict the target features for arbitrary DNA proxies in the next lineage, including other data points not in the tribe of Cichlidae for the problem. This section explains how these models were obtained and provides an assessment as to how they might answer the major question whether evolution in a lineage is entirely random.

Predicting Divergence Time (L, f)

Firstly, the NN model used to solve this problem was developed by pre-processing the DNA sequences obtained for the two lineages as described in Section 2.2, followed by initializing the network architecture and training using backpropagation. During preprocessing, nxh bases 4mP3 and 5mP6

were used to convert the DNA sequences of the NADH and COI genes for the Cichlidae (C) lineage into numerical vectors of length 3 and 6 respectively. Similarly, nxh bases 4mP3 and 5miC3Mg were used for the HelicoBacter dataset to convert the NixA+trpC genes into numerical vectors of length 3 and 3 respectively. The dataset was split as 70%/30% for training/ testing but sometimes the split had to be adjusted to 80%/20% to obtain models with better predictions. The target label for this problem is the divergence time for each organism.

Since the target labels are continuous quantities (time) the metric for evaluation of the quality of the NN solutions is usually chosen to be the average percentage in the relative error across, given by

$$\mathrm{RE} = \frac{|actual - predicted|}{actual} * 100$$

for the predictions as a consolidated measure of how well the NN performed across all predictions in the testing phase. The results are shown in **Table 4**, together with the choices for preprocessing and the NN models.

Lineage	***nxh base***	***Arch. (hidden layers)***	***Min / Avg /Max RE (%)***
Cichlidae Fish - NADH	4mP3	[2]	0 / 31 / 43.64
	4mP3+5mP36	[9,6,3,1]	0 / 7 / 70
	5mP36	[6,5,1]	0 / 19 / 52
Cichlidae Fish - COI	4mP3	[2]	0 / 7 / 22
	5mP36	[6,3,2,1]	0 / 5 / 41
	4mP3+5mP36	[9,6,3,1]	0 / 7 / 12
HelicoBacter-NixA+trpC	4mP3	[2]	0 / 27 / 46
	5miC3Mg	[2]	0 / 28 / 47

Table 4.
Relative Error (RE) of the solutions for Problem DIVTIME.
The best prediction is given by the COI proxy for lineage C on the concatenation of signatures on 4mP3 and 5mP3 nxh bases, while 4mP3 performs unsatisfactorily for both C and H.

Predicting Difference in Divergence Time (L, f)

The NN models to solve this problem were obtained using the same process described in Section 3.1. The only difference was that a different architecture was chosen to better fit the input features for the two proxies used to predict the difference in divergence times. The feature vectors were obtained by concatenating the two feature vectors of two different input

organisms from the datasets, providing the neural network with enough information to make predictions. The target label was the difference in their divergence times. **Table 5** summarizes the results.

Lineage	***nxh base***	***Arch. hidden layers)***	***Min / Avg / Max RE (%)***
Cichlidae Fish - NADH	4mP3	[8]	0 / 8 / 34 *
Cichlidae Fish - COI	4mP3	[3]	0 / 6 / 10 *
HelicoBacter-NixA+trpC	4mP3	[3]	0 / 50 / 80 *

Table 5.
Relative Error (RE) of the solutions for Problem [DDIV].
The best prediction is given by the COI proxy for lineage C on the signatures of nxh basis 4mP3, while the same nxh basis performs inadequately for H.

*Some large value outliers were excluded in the RE calculation due to the significant difference in divergence time between the newer and older species, which led to a high RE value from the original divergence.

Predicting Next Proxy (L, f)

Likewise, the feature vectors for this problem are obtained by concatenating the feature vectors of the organism and its three ancestors in the lineage. Since the first three organisms do not have three ancestors, these datapoints were ignored. Predicting the actual DNA sequence of the next organism in the lineage is a very difficult problem because it involves many other factors (e.g., environmental) than just genetic inheritance obtained from its ancestors. Thus, in this problem, the aim was reduced to predicting the genomic signature of the next organism than the entire DNA sequence. To enable the predictions, the 3D feature vectors (x,y,z), from the nxh bases 4mP3 and 5miC3Mg, were reduced to a 2D feature vector (x',y') by a geometric transformation (rotations and translations). This is possible since the feature vectors are normalized, i.e., lie on a plane in 3D Euclidean space given by the condition x+y+z=1. These transformations adjusted the orientation and position of the points but only require two coordinates. Thus, the 3D vectors are reduced to 2D vectors while hopefully retaining the underlying information in the original genomic signatures. As a result, the NN architecture will require only two output neurons for the two features in the genomic signatures being predicted. After this, the learning takes place as explained earlier. Once the required accuracy

threshold is reached during training, the overall relative error for the predictions were calculated. The results are summarized in **Table 6**.

Lineage	nxh base	*Arch. hidden layers)*	*Min / Avg / Max RE (%)*
Cichlidae Fish - NADH	4mP3	[3,4]	0 / 189 / 520
HelicoBacter- NixA+trpC	4mP3	[2]	0 / 95 / 500
HelicoBacter -NixA+trpC	5miC3Mg	[2]	0 / 57 / 900

Table 6.
Relative Error (RE) of the solutions for Problem [NEXTP].
The best prediction is given by the NixA+trpC proxies for lineage H on the signatures of nxh basis 5miC3Mg , while the signatures of the nxh basis 4mP3 performs unsatisfactorily for C.

To assess the significance of these results for the three problems, one must consider a few factors. First, the difficulty of the problems, as discussed in their statements above. Second, when the DNA sequences were transformed into numerical vectors through the dimensionality reduction process, some crucial information must have been inevitably lost and therefore impacted the models' ability to predict the labels. Nonetheless, the neural networks were indeed able to make the most of the information given to them and make predictions out of that data, albeit imperfectly and thus, the DNA must contain some information for the NN to extract the target features. On the other hand, this fact raises deeper questions about the reason for the errors: Is the loss of information just in the pre-processing (from DNA to genomic signatures.) Does evolution lose some information in the way it is encoding the divergence time for example?

Secondly, it is important to evaluate the magnitude of the relative errors (REs) and their significance for the research question. In the problem [DIVTIME], the lowest relative error was 5%. In terms of millions (M) of years, a 5% error is just 50,000 years. Given the scale of evolutionary time, this level of error is therefore acceptable, and this NN could be used for such datasets producing predictions of divergence time with a high predictability. Similarly, for the second problem [DDIV], a 6% average RE (about 60,000 years) is a relatively close prediction of the difference in the divergence times at a time scales of 36M years. Even the traditional methods used to do time-series analysis using calibrated radiocarbon dates, which often have irregular uncertainties (Carleton, 2018). Finally, problem [NEXTP] was the most challenging to solve, as the NN had to

forecast the genomic signatures for the next species with a minimum RE of 57% is surprisingly good, but a maximum RE of 189% is not surprising. Alternatively, it is possible that the target labels are themselves in error with respect to the gold standard of the historical record.

All things considered, it is evident that there is substantial evidence in these results to answer the question of whether evolution is entirely random. The fact that the neural network was able to predict the divergence times and the difference in the divergence time with a high degree of correctness despite a number of possibly faulty assumptions, implies that there are underlying patterns and predictability in the evolutionary process. While randomness likely plays a role in evolution, this unpredictability could stem from genetic mutations that occur randomly or from inherent limitations in the data. But the ability of such models to make informed predictions on some of the events is strong evidence that deterministic factors are contributing to shaping evolution. This challenges the assumption that evolutionary events are purely random and points to the possibility of them being a more organized process influenced by both random and predictable elements. The results of this paper also raise a deeper question: Are there random events in the evolutionary process that are provably inherently unpredictable?

Discussion and Conclusion

This paper has addressed a fundamental research question whether evolutionary events are entirely random, following up on recent studies that it might not be. The working definition of randomness was provided in the introduction as a sequence of numbers which contains no recognizable patterns or regularities that it is impossible to predict the next number in the sequence either by humans or using a computer program. The main takeaway is that the results provide quantitative evidence indicating that evolutionary events involve predictable elements along with random processes. Neural network's abilities to provide an approximation, with low relative error, of the divergence time of organisms or the (incomplete but significant) information about the genetic composition of their descendants in a lineage, i.e., genomic signatures of the DNA, suggests that there are indeed deterministic factors in evolutionary processes. Since the Min RE= 0 in all three problems, there were instances of individuals in each where the neural network made perfect predictions that matched the given label, although there were also individuals for which RE > 0 because the average RE was positive. This would appear to be in direct opposition to the conventional view of evolutionary events being essentially random.

This study could have used a larger and more representative dataset. But due to difficulties obtaining reliable data, the data set used in this study is small and narrowly focused. Nonetheless, this restriction is a necessary first step to set up the stage for a more comprehensive future study aimed at elucidating the relative contribution of environmental factors to the randomness in evolutionary processes.

On the other hand, acknowledging the limitations is necessary for these results. First, this study only included two lineages, which limits the generalizability of the results to a broader range of evolutionary scenarios. Second, a small or narrowly focused dataset may not contain enough variability to fully represent genetic diversity present in larger divergent groups, potentially leading to overfitting or missed patterns in the models. Third, it is clear in biology that evolution is influenced by both genetic and environmental factors, but the models were trained using only DNA sequences without attempting to incorporate additional information about the environment (largely unknown at divergence times), which could have provided a more comprehensive understanding of evolutionary patterns, and the effect of variables such as climate and geography.

Despite these limitations, the quantitative analyses of the predictions through low relative errors suggest a significant degree of predictability for these two lineages. Based on the results, it can be confidently stated that evolutionary events contain predictable features. Of course, that does not mean that this evidence can conclude that evolutionary events are entirely predictable. Since the study was conducted on only two lineages, while many others remain untested, the findings cannot be generalized to arbitrary lineages, although the methods are based on the deep structure of DNA that is pervasive through the entire biome. In some sense, this study really raises more questions than it answers, but these questions about predictability are of enormous interest given the uncertain future of many species due to the increasing fragility of our planet (e.g. due to climate changes.) Much more comprehensive and difficult studies are needed to confirm that evolution is not entirely random across the full range of evolutionary events. For example, can we ascertain the existence of a specific evolutionary event and strong evidence that it is random?

Acknowledgment

We would like to acknowledge Drs. Luis Fernando Garcia and Alexander Colorado of the National U of Colombia for their invaluable assistance in suggesting appropriate datasets, the specific genes and DNA sequences

and collecting and furnishing the divergence times for the various organisms. Their support was essential to the success of this study.

References

1. Paley, W. (1802). Natural theology: Or, Evidence of the existence and attributes of the deity, collected from the appearances of nature (2nd ed.). R Faulder. https://doi.org/10.1037/11747-000

2. Darwin, Charles (1859). On the Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life. London: John Murray, 1859.

3. Dobzhansky, T. (1950). Mendelian populations and their evolution. The American Naturalist,84:819, 401—418.

4. Wortel MT, Agashe D, Bailey SF, Bank C, Bisschop K, Blankers T, Cairns J, Colizzi ES, Cusseddu D, Desai MM, van Dijk B, Egas M, Ellers J, Groot AT, Heckel DG, Johnson ML, Kraaijeveld K, Krug J, Laan L, Lässig M, Lind PA, Meijer J, Noble LM, Okasha S, Rainey PB, Rozen DE, Shitut S, Tans SJ, Tenaillon O, Teotónio H, de Visser JAGM, Visser ME, Vroomans RMA, Werner GDA, Wertheim B, Pennings PS. Towards evolutionary predictions: Current promises and challenges. Evol Appl. 2022 Dec 9;16(1):3-21. doi: 10.1111/eva.13513. PMID: 36699126; PMCID: PMC9850016.

5. Mas A, Lagadeuc Y, Vandenkoornhuyse P. Reflections on the Predictability of Evolution: Toward a Conceptual Framework. iScience. 2020 Oct 27;23(11):101736. doi: 10.1016/j.isci.2020.101736. PMID: 33225244; PMCID: PMC7666346.

6. Garzon, M., Yang, C., Venugopal, D., Kumar, N., Jana, K., & Deng, L. (2022). Dimensionality Reduction in Data Science. 1-265. Springer Nature. https://doi.org/10.1007/978-3-031-05371-9

7. Carleton WC, Campbell D, Collard M. Radiocarbon dating uncertainty and the reliability of the PEWMA method of time-series analysis for research on long-term human-environment interaction. PLoS One. 2018 Jan 19;13(1):e0191055. doi: 10.1371/journal.pone.0191055. PMID: 29351329; PMCID: PMC5774753.

8. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20-D26. doi: 10.1093/nar/gkab1112. PMID: 34850941; PMCID: PMC8728269.

9. Wikipedia contributors. (2025, January 1). Statistical randomness. Wikipedia. https://en.wikipedia.org/wiki/Statistical_randomness

Jacob Stewart is a Physics and Mathematics double major studying at the University of Memphis, graduating in May 2026. He has done research in mathematics, physics, and mechanical engineering. He has done work with Dr. Thang Hoang, Dr. Thomas Hagen, and Dr. Ranganathan Gopalakrishnan. He currently works in plasma physics and hopes to attend graduate school for aerospace studies in fall 2027. The research for *QuaesitUM* was funded through a summer research fellowship from the Helen Hardin Honors College. Jacob recently presented this work at the National Conference for Undergraduate Research and hopes to continue work on undergraduate research on his path towards graduation.

Jacob Stewart

Escape and Pursuit of an L^1 Pursuer and an L^2 Target

Faculty Sponsor
Dr. Thomas Hagen

Abstract

In mathematical pursuit and escape games, a pursuer (agent 1) tries to catch a target (agent 2) by closing the distance between them. Both pursuer and target can move as long as their movement is unobstructed. The effects of limiting one of these agents to movement measured in the L^1 metric and the addition of a finite, straight-line barrier were investigated. It was found that even with lower speed (up to a ratio of $\frac{1}{\sqrt{2}}$), an L^2 target can escape from an L^1 pursuer if the target takes the "optimal" path. When introducing the barrier, there are three cases for the L^1 pursuer's path: one axis of motion (AoM) blocked, both AoM blocked, and no AoM blocked. The cases with single AoM and both AoM being blocked results in global change of the dominance regions, whereas the case with no AoM blocked only results in local change.

Introduction

Pursuit and Escape (PE) games are used to define the behavior of two or more agents in which one or more pursuer is attempting to capture one or more target (Nahin 1981). These games are usually done in a two-dimensional case to simplify modeling. There are a variety of ways to describe the way in which two agents participating in a game interact with one another, but this paper describes interactions between agents whose movement are described with respect to two different norms.

DEFINITION 1 (L^1 and L^2 Norm). *Let $P = (x, y)$ be a point in $\mathbb{R}^2$. Then,*

$$(1)\ ||P||_1 = ||(x,y)||_1 = |x| + |y| \quad \textit{is the } L^1 \textit{ norm of } (x,y)$$

$$(2)\ ||P||_2 = ||(x,y)||_2 = \sqrt{x^2 + y^2} \textit{ is the } L^2 \textit{ norm of } (x,y)$$

DEFINITION 2 (L^1 and L^2 Metric). *Let $P = (x_1, y_1)$, $Q = (x_2, y_2)$ be points in $\mathbb{R}^2$. Then,*

(1) The L^1 metric of P and Q is given by $d_1(P,Q) = ||P - Q||_1 = ||(x_1 - x_2, y_1 - y_2)||_1$

(2) The L^2 metric of P and Q is given by $d_2(P,Q) = ||P - Q||_2 = ||(x_1 - x_2, y_1 - y_2)||_2$

A metric is also called a distance because it measures how far apart two points are. The L^2 metric is commonly referred to as the Euclidean metric or distance.

DEFINITION 3 (Apollonian Circle). *For two distinct points A and B in $\mathbb{R}^2$, the set of all points $P = (x, y)$ such that, for some constant $c > 0$,*

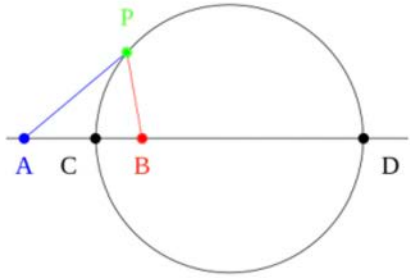
$$d_{1,2}(P, A) = c d_{1,2}(P, B)$$

is called an Apollonian Circle.

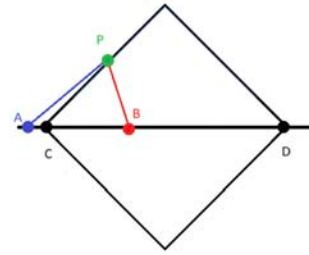
REMARK. *Apollonian Circles are circles in the classical sense, as illustrated in Figure 1.1a.*

REMARK. *Changing to the L^1 metric, i.e $d = d_1$, results in diamond shaped Apollonian 'circles.' As illustrated in Figure 1.1b.*

This paper has two objectives: 1) analyzing the movement of two agents, one being limited to movement measured with respect to the L^1 metric and the other with respect to the L^2 metric and 2) illustrating the effect of a finite, straight line barrier on the agents' movement and interactions.



(A) Apollonian circle definition for the L^2 metric in which P represents the point that has a ratio constant for $c = \frac{d_2(P,A)}{d_2(P,B)}$.



(B) Apollonian circle definition for the L^2 metric in which P represents the point that has a ratio constant for $c = \frac{d_1(P,A)}{d_1(P,B)}$.

Figure 1.1 Apollonian Circle in L^2 and L^1 Metric

These objectives are motivated by certain predator-prey interactions, as well as military applications like unmanned ground vehicle (UGV) to unmanned aerial vehicle (UAV) pursuit and target-tracking algorithms. The most interesting cases arise with an agent on the ground in a city with a grid-like structure, representing the L^1 metric case, and an aerial agent that is free to move above the city, representing the L^2 metric case.

The literature covers a substantial amount of work on traditional escape and pursuit games, particularly involving single target, single pursuer games, shown in (Nahin 1981). This is not necessarily new, but a paper by (Oyler et al. 2016) expands upon previous works to add complexity to the problem of barrier addition and how dominance regions behave around said barrier among other contributions that are less relevant to this paper.

NOTE. For the duration of this paper, we consider two agents: Agent A (the pursuer) and Agent B (the target). Agent A is restricted to movement measured with respect to the L^1 metric ($d_A = d_1$) with starting point (x_A, y_A) and constant speed $v_A > 0$. Agent B follows for the L^2 metric ($d_B = d_2$) with starting point (x_B, y_B) and constant speed $v_B > 0$.

DEFINITION 4 (Isochronic Curve). *Let $\mathcal{U}$ be a region in $\mathbb{R}^2$ and $(x_o, y_o) \in \mathcal{U}$. An Isochronic Curve in $\mathcal{U}$ is a continuous curve $(x(t), y(t)) : [0, T] \to \mathbb{R}^2$, such that*

$$\begin{cases} (x(0), y(0)) = (x_o, y_o) & \\ d((x(t), y(t)), (x_o, y_o)) \leq vt & \text{for } 0 \leq t \leq T \\ (x(t), y(t)) \in \mathcal{U} & \text{for } 0 \leq t \leq T \end{cases} \tag{1.1}$$

where d is either the L^1 or L^2 metric. If equality holds above in (1.1)*, the corresponding isochronic curve is called an isochrone.*

DEFINITION 5 (Isochronic Set). *Let $\mathcal{U}$ be a region in $\mathbb{R}^2$, $(x_o, y_o) \in \mathcal{U}$, and $v > 0$. For $T > 0$, the set*

$$I(T) = \{(x, y) \in \mathcal{U} \mid \exists \text{ an isochronic curve in } \mathcal{U}, (x(t), y(t)), 0 \leq t \leq T \text{ such that } (x(0), y(0)) = (x_o, y_o), (x(T), y(T)) = (x, y)\}$$

The set $I(T)$ is called the Isochronic Set at time T.

For Agent A, we consider $I(T)$ with $d = d_A$, $v = v_A$, $(x_o, y_o) = (x_A, y_A)$ and write $I_A(T)$. For Agent B, we consider $I(T)$ with $d = d_B$, $v = v_B$, $(x_o, y_o) = (x_B, y_B)$ and write $I_B(T)$.

REMARK. *If $\mathcal{U} = \mathbb{R}^2$ the set $I(T)$ reduces to*

$$I(T) = \{(x, y) \in \mathcal{U} \mid d((x, y), (x_o, y_o) \leq vt, 0 \leq t \leq T\}$$

This will define circles if $d = d_2$ and diamond shapes if $d = d_1$.

Dominance Regions are areas over which an agent is able to reach any point in that are before another agent.

DEFINITION 6 (Dominance Region). *The Dominance Region of Agent B is defined as*

$$\bigcup_{t^* > 0} I_B(t^*) \backslash I_A(t^*)$$

REMARK. *We interpret an unbounded Dominance Region for Agent B to mean that Agent B can always reach points at any time $t^* > 0$ which cannot be reached by Agent A for the same time, t^*.*

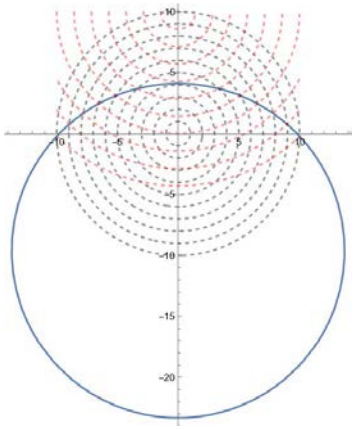


Figure 1.2
Agent one (located at the origin) has a bound dominance region because the intersection of the two agent's isochrones (dotted lines) defines a closed circle.

EXAMPLE 1. **Figure 1.2** *shows two agents whose movement are described with respect to the L^2 metric on a two dimensional plane, one positioned at the origin, and the other ten units away along the y-axis. The agent starting at the origin is given a lower starting speed than the other. This results in its dominance region being bounded because the agent is unable to escape the other. This is because they both have their movement measured with respect to the same metric.*

NOTE. *This paper explores the changes when one agent's movement is measured with respect to the L^2 metric and the other's is measured with respect to the L^1 metric. This gives the agent moving with respect to the L^2 metric an unbounded dominance region even if it is slower than the other agent, up to a later defined velocity ratio.*

This leads to the main result of this work:

THEOREM (L^1-L^2 Norm Interaction Theorem). *Suppose Agent A has starting point (c, d), for $c \geq 0$ and $d > 0$, and Agent B has starting point $(0, 0)$. For a velocity ratio, $\gamma = \frac{v_A}{v_B} \geq \frac{1}{\sqrt{2}}$, then the Dominance Region of Agent B, defined by $\bigcup_{t^*>0} I_B(t^*) \backslash I_A(t^*)$, is unbounded.*

The proof of this theorem is given in the following section. The last section of this work will consider the impact of a single, finite straight-line barrier on the Dominance Regions of both agents through numerical investigation. A rigorous statement regarding this is left for future work.

Theorem Background and Proof

This project started through modeling the movement of two agents using isochrones to define their dominance regions. After limiting one agent to movement measured with respect to the L^1 metric, the equations for these isochrones were de-

fined and used to describe new dominance regions. Then, the asymptotic behavior of the curves defining the dominance region was identified, proving that the target can escape for certain 'velocities' lower than the pursuer's velocity.

To start, some results from Oyler's paper were recreated with both agents having their movement measured with respect to the L^2, or Euclidean, norm (Oyler et al. 2016). Desmos was used to plot the initial graphics for both agents. Later, Wolfram Mathematica was used to plot the more complex cases. Note that all cases were computed with $\gamma = \frac{v_A}{v_B} = \sqrt{0.6}$ and a separation of $d = 5$ for purposes of illustration. γ is the ratio of velocities between the two agents, and d is the separation between them along the y-axis.

Movement Measured with respect to the L^2 Metric

With both agent's movement being measured with respect to the L^2 metric, the distance traveled by an agent is defined using Cartesian coordinates as $x^2 + y^2 = v^2t^2$, with x and y being the coordinates of the agent and v being the velocity of the agent. The previously mentioned isochrone corresponds to a circle for every value of t, meaning that it would take the agent the same amount of time, t, to reach any point on that circle, provided it starts from the center of the circle. To model the dominance region, the intersection of the agent's isochrones are found. To do this, the left hand sides of both agent's movement equations are set equal to one another.

Both agents are each given different starting points and speeds, then the time component is removed because their movement can be defined entirely with these movement functions and the resulting isochrones. First, we define the isochrone equations of Agent A and Agent B: Agent A, $x_A^2 + y_A^2 = v_A^2t^2$ and Agent B, $x_B^2 + y_B^2 = v_B^2t^2$. Agent A's starting location, (x_A, y_A), should be different from Agent B's, (x_B, y_B). We start Agent B at the origin and Agent B at (c, d) with $c \geq 0$ and $d > 0$. We let $\gamma = \frac{v_B}{v_A}$, $x = x_B = x_A$, and $y = y_B = y_A$, then this simplifies to

$$x^2 + y^2 = \gamma((x - c)^2 + (y - d)^2) \tag{2.1}$$

Figure 1.2 shows how the isochrones of each agent intersect to from a circle. The solid line circle is enclosed, meaning that Agent B is not able to escape. This defines a bounded dominance region.

Limitation to the L^1 Norm

One agent is limited to movement measured in the L^1 norm to study its interactions with another agent that has its movement measured with respect to the L^2 norm. Limiting one agent to movement measured with respect to the L^1 norm results in the corresponding equation denoting maximal distance traveled in time t becoming $|x - c| + |y - d| = vt$. In this case, the isochrones will change their shape from being circles to being diamond shaped.

We concentrate on the case where $\mathcal{U} = \mathbb{R}^2$. $I_B(t^*)$ has as its boundary all points (x, y) such that $x^2 + y^2 = v_B^2t^{*2}$, and $I_A(t^*)$ has as its boundary all points (x, y) such that $(|x - c| + |y - d|)^2 = v_A^2t^{*2}$.

Hence, to understand the Dominance Region, we focus on the point (x, y) that are shared by $I_A(t^*)$ and $I_B(t^*)$, i.e. (x, y) such that

$$x^2 + y^2 = \gamma^2((x - c)^2 + 2|x + c||y - d| + (y - d)^2) \tag{2.2}$$

This equation plots the shared boundary of $I_A(t)$ and $I_B(t)$ for any $t > 0$, and, depending on the value of gamma, will lead to a bounded or unbounded dominance region for Agent B.

Proof for Theorem

L^1-L^2 Norm Interaction Theorem

Suppose Agent A has starting point (c, d), for $c \geq 0$ and $d > 0$, and Agent B has starting point $(0, 0)$. For a velocity ratio, $\gamma = \frac{v_A}{v_B} \geq \frac{1}{\sqrt{2}}$, then the Dominance Region of Agent B, defined by $\bigcup_{t^*>0} I_B(t^*) \backslash I_A(t^*)$, is unbounded.

NOTE. This is because movement measured with respect to the L^1 metric is symmetric only along the axes, i.e. through a reflection over either axis or a rotation of $90°$. It is therefore sufficient to only consider cases where $c \geq 0$ and $d > 0$. We divide the proof for this theorem into two cases: $c = 0$ and $c > 0$. Starting with $c = 0$:

PROOF.

NOTE. This proof can be divided into four cases: $\gamma > 1$, $\gamma = 1$, $\frac{1}{\sqrt{2}} < \gamma < 1$, and $\gamma = \frac{1}{\sqrt{2}}$. When referring to the "region," the dominance region of Agent B (agent starting at the origin that is orange) is being described. Also, recall that γ is the velocity of Agent B (orange) divided by the velocity of Agent A (blue).

For the case where $\gamma > 1$, it is obvious that Agent B can escape because it is able to move faster that Agent A, thus rendering capture impossible. Thus, the region is unbounded for $\gamma > 1$.

For $\frac{1}{\sqrt{2}} < \gamma < 1$, we postulate the equations for the boundary of the dominance regions are approximated by oblique asymptotes of form $y = mx + b$ determined by equation $x^2 + (ax + b)^2 = \gamma^2(x^2 + 2|x||ax + b - d| + (ax + b - d)^2) + \varepsilon(x)$, with $\varepsilon(x)$ being an error term. As $x \longrightarrow \infty$ where $\varepsilon(x) \longrightarrow 0$ as $x \longrightarrow \infty$.

There are four equations determined for the range $\frac{1}{\sqrt{2}} < \gamma < 1$, with the quantity a being $\pm\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2}$ and $\pm\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2}$ and the quantity b being $\frac{d\gamma^2(1-a)}{a+\gamma^2(1-a)}$.

These four equations are as follows:

$$
\begin{cases}
(1)\ y = \dfrac{\gamma^2 + \sqrt{2\gamma^2 - 1}}{1 - \gamma^2}x + \dfrac{d\gamma^2(1 - \frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2})}{\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(1 - \frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2})} \\
(2)\ y = \dfrac{\gamma^2 - \sqrt{2\gamma^2 - 1}}{1 - \gamma^2}x + \dfrac{d\gamma^2(1 - \frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2})}{\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(1 - \frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2})} \\
(3)\ y = -\dfrac{\gamma^2 + \sqrt{2\gamma^2 - 1}}{1 - \gamma^2}x + \dfrac{d\gamma^2(1 + \frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2})}{-\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(1 + \frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2})} \\
(4)\ y = -\dfrac{\gamma^2 - \sqrt{2\gamma^2 - 1}}{1 - \gamma^2}x + \dfrac{d\gamma^2(1 + \frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2})}{-\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(1 + \frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2})}
\end{cases}
$$

With these oblique asymptotes in place, it becomes clear the region of dominance is unbounded. By definition, these asymptotes will not intersect outside of the starting region. Thus, Agent B is able to escape along this open region for the range $\frac{1}{\sqrt{2}} < \gamma < 1$.

The equations for $\gamma = \frac{1}{\sqrt{2}}$ and $\gamma = 1$ still remain. The case $\gamma = 1$ is examined next.

For $\gamma = 1$, simplify equation 2.2: $2|x||y - d| - 2dy + d^2 = 0$.

Assume $|x|$ is very large.

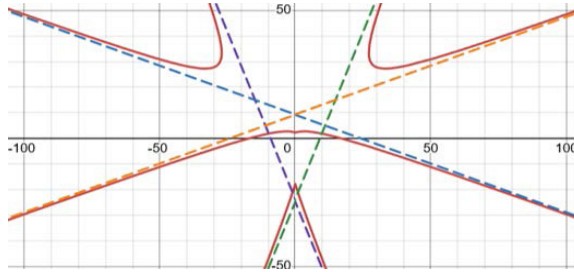


Figure 2.1
Oblique Asymptotes for $\gamma = \sqrt{0.6}$. (1): Green, (2): Orange, (3): Purple, (4): Blue

Assuming $x \gg 0$ and $y \geq d$, the equation becomes $-2x(d - y) - 2dy + d^2 = 0 \implies y(2x - 2d) = 2xd - d^2 \implies y = d\frac{2x-d}{2x-2d}$, and, as $x \longrightarrow \infty$, $y \longrightarrow d$.

Assuming $x \ll 0$ and $y \geq d$, the equation becomes $-2x(y - d) - 2dy + d^2 = 0 \implies y(2x + 2d) = 2xd - d^2 \implies y = d\frac{2x-d}{2x+2d}$, and, as $x \longrightarrow -\infty$, $y \longrightarrow d$.

Assuming $x \gg 0$ and $y < d$ results in $y = d\frac{2x+d}{2x+2d} = d\frac{2+\frac{d}{x}}{2+\frac{d}{x}}$. As $x \longrightarrow \infty$, $y \longrightarrow d$.

Assuming $x \ll 0$ and $y < d$ results in $y = d\frac{2x-d}{2x-2d}$. As $x \longrightarrow -\infty$, $y \longrightarrow d$.

This results in a single horizontal asymptote at $y = d$.

Now, assume y is very large.

Assuming $y \gg 0$ and $x > d$ results in $2x = \frac{2dy+d^2}{y-d}$. As $y \longrightarrow \infty$, $x \longrightarrow d$.

Assuming $y \ll 0$ and $x < -d$ results in $-2x = \frac{2dy+d^2}{y-d} = -d$. As $y \longrightarrow \infty$, $x \longrightarrow -d$.

Assuming $y \ll 0$ and $x > 0$ results in $2x = \frac{2dy+d^2}{y-d}$. $y \longrightarrow \infty$, $x \longrightarrow -d$. This is a contradiction.

Assuming $y \gg 0$ and $x < 0$ results in $2x = \frac{2dy+d^2}{y-d}$. As $y \longrightarrow \infty$, $x \longrightarrow d$. This is a contradiction.

This corresponds to two vertical asymptotes at $x = \pm d$. Therefore, Agent B's dominance region is not bounded for $\gamma = 1$. The results are plotted in Figure 2.5.

The only remaining case is $\gamma = \frac{1}{\sqrt{2}}$. The plot of the equation for this γ value will result in four parabolic branches.

Because of the term $|x|$, through symmetry, the solution for $x \geq 0$ is sufficient for both cases.

Simplifying yields the following equation:

$x^2 + y^2 - 2x|y - d| + 2dy - d^2 = 0$

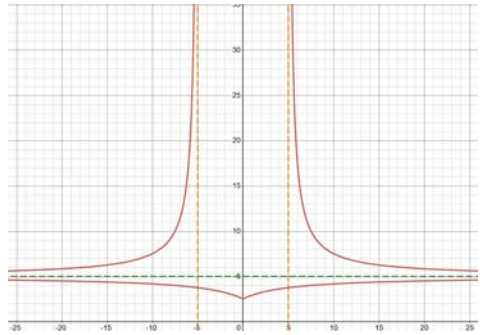


Figure 2.2

Plot for $\gamma = 1$ and $d = 5$ with horizontal and vertical asymptotes

We solve for $y \leq d$ because we are interested in the Dominance Region of Agent B:

$x^2 + y^2 - 2dy + 2dx + 2dy - d^2 = 0. \implies x^2 + y^2 - 2xd + 2xy + 2dy - d^2 \implies$
$y^2 + 2(d + x)y + x^2 - 2xd - d^2 = 0$

This is a quadratic. Solving for $y_{1,2}$:

$$y_{1,2} = \frac{-2(d+x) \pm \sqrt{4(x+d)^2 - 4(x^2 - 2xd - d^2)}}{2} = -(x + d) \pm \sqrt{x^2 + 2xd + d^2 - x^2 + 2xd + d^2}$$

$$y_{1,2} = -(x + d) \pm \sqrt{4xd + 2d^2}$$

These solutions for y define parabolic branches that describe the boundary of the Dominance Region of Agent B for $x \geq 0$ and $y \leq d$. It is therefore clear that this Dominance Region is unbounded. See Figure 2.3.

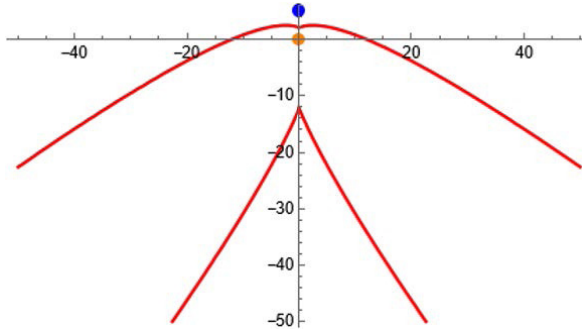


Figure 2.3

Parabolic Branches for $\gamma = \frac{1}{\sqrt{2}}$ and d $= 5$

Each of the proposed four cases have been proven to be unbounded regions for Agent B, thus, Agent B's dominance region is unbounded for $\gamma \geq \frac{1}{\sqrt{2}}$.□

This was the case where Agent A starts at $(0, d)$ and Agent B starting at $(0, 0)$. If Agent A is instead started from (c, d), with $c \neq 0$, the proof will differ in every case. We now take $c > 0$:

PROOF.

NOTE. We take the same four cases for this proof as for the above proof.

For the case where $\gamma > 1$, it is obvious that Agent B can escape because it is able to move faster that Agent A, thus rendering capture impossible. Thus, the region is unbounded for $\gamma > 1$.

For $\frac{1}{\sqrt{2}} < \gamma < 1$, we postulate the equations for the boundary of the dominance regions are approximated by oblique asymptotes of form $y = mx + b$ determined by equation $x^2 + (ax + b)^2 = \gamma^2((x - c)^2 + 2|x - c||ax + b - d| + (ax + b - d)^2) + \varepsilon(x)$, with $\varepsilon(x)$ being an error term. As $x \longrightarrow \infty$, where $\varepsilon(x) \longrightarrow 0$ as $x \longrightarrow \infty$. The slope values, a, will be the same as before because a change in c will only shift the isochrones of Agent A. Only the value of b will change to include the new c term.

The quantity a is $\pm\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2}$ or $\pm\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2}$. The quantity b is now $\frac{d\gamma^2(1-a)}{a+\gamma^2(c+1-a(1+c))}$.

These four equations are as follows:

$$\begin{cases} (1)\ y = \dfrac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2}x + \dfrac{d\gamma^2(1-\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2})}{\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(c+1-\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2}(c+1))} \\ (2)\ y = \dfrac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2}x + \dfrac{d\gamma^2(1-\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2})}{\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(c+1-\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2}(c+1))} \\ (3)\ y = -\dfrac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2}x + \dfrac{d\gamma^2(1+\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2})}{-\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(c+1+\frac{\gamma^2+\sqrt{2\gamma^2-1}}{1-\gamma^2}(c+1))} \\ (4)\ y = -\dfrac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2}x + \dfrac{d\gamma^2(1+\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2})}{-\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2} + \gamma^2(c+1+\frac{\gamma^2-\sqrt{2\gamma^2-1}}{1-\gamma^2}(c+1))} \end{cases}$$

With these oblique asymptotes in place, it becomes clear the region of dominance is unbounded. Thus, Agent B is able to escape along this open region for the range $\frac{1}{\sqrt{2}} < \gamma < 1$.

The equations for $\gamma = \frac{1}{\sqrt{2}}$ and $\gamma = 1$ still remain. The case $\gamma = 1$ is examined next.

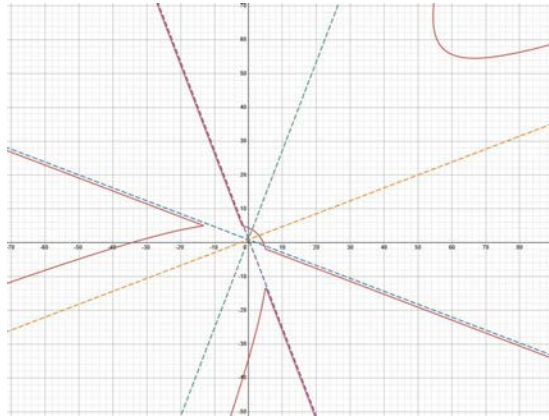


Figure 2.4
Oblique Asymptotes for $\gamma = \sqrt{0.6}$ and Agent A starting location at (5,5).
(1): Green, (2): Orange, (3): Purple, (4): Blue

For $\gamma = 1$, simplify equation 2.2: $-2cx + c^2 + 2|x - c||y - d| - 2dy + d^2 = 0$.

Assume $|x|$ is very large.

Assuming $x \gg 0$ and $y \geq d$, the equation becomes $-2x(d-y) - 2cx - 2y(c+d) + c^2 + d^2 = 0 \implies y(2x - 2c - 2d) = 2x(c+d) - c^2 - d^2 \implies y = \frac{2x(c+d)-c^2-d^2}{2x-2c-2d}$, and, as $x \longrightarrow \infty$, $y \longrightarrow c + d$.

Assuming $x \ll 0$ and $y \geq d$, the equation becomes $-2x(y-d) - 2cx - 2y(c+d) - c^2 + d^2 = 0 \implies y(-2x - 2c - 2d) = 2x(c-d) - c^2 - d^2 \implies y = -\frac{2x(c-d)-c^2-d^2}{2x+2c+2d}$, and, as $x \longrightarrow -\infty$, $y \longrightarrow d - c$. This is a contradiction for $c \neq 0$

Assuming $x \gg 0$ and $y \leq d$ results in $y = \frac{2x(c+d)-c^2-d^2}{2x-2d}$. As $x \longrightarrow \infty$, $y \longrightarrow d - c$.

Assuming $x \ll 0$ and $y \leq d$ results in $y = \frac{2x(c+d)-c^2-d^2}{2x-2d}$. As $x \longrightarrow -\infty$, $y \longrightarrow c + d$. This is a contradiction for $c \neq 0$

This results in two horizontal asymptotes at $y = c + d$ and $y = d - c$.

Now, assume y is very large.

Assuming $y \gg 0$ and $x > c$ results in $x = \frac{2y(c+d)-c^2-d^2}{2y-2c-2d}$. As $y \longrightarrow \infty$, $x \longrightarrow c + d$.

Assuming $y \ll 0$ and $x > c$ results in $x = -\frac{2y(-c+d)-c^2-d^2}{2y+2c+2d}$. As $y \longrightarrow \infty$, $x \longrightarrow c - d$. This is a contradiction.

Assuming $y \gg 0$ and $x < c$ results in $x = \frac{2y(-c+d)-c^2-d^2}{2y+2c+2d}$. $y \longrightarrow \infty$, $x \longrightarrow c - d$.

Assuming $y \gg 0$ and $x < c$ results in $x = \frac{2y(c+d)-c^2-d^2}{2y-2c-2d}$. As $y \longrightarrow \infty$, $x \longrightarrow c + d$. This is a contradiction.

This corresponds to two vertical asymptotes at $x = c + d$ and $x = c - d$. Therefore, Agent B's dominance region is not bounded for $\gamma = 1$. The results are plotted in Figure 2.5.

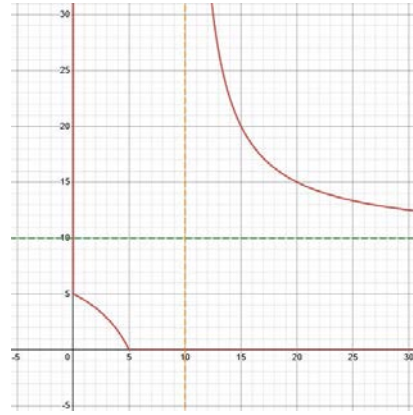


Figure 2.5
Plot for γ = 1 and c = d = 5 with horizontal and vertical asymptotes

The only remaining case is $\gamma = \frac{1}{\sqrt{2}}$. The plot of the equation for this γ value will result in at most four parabolic branches.

Simplifying yields the following equation:

$$x^2 + y^2 - 2|x - c||y - d| + 2cx + 2dy - c^2 - d^2 = 0$$

Assume $y \leq d$ because the Dominance Region of Agent B is of interest. We start with the case with $x > c$:

$$x^2 + y^2 + 2xy - 2dx + 2cd - 2cy + 2cx + 2dx + 2dy - c^2 - d^2 = 0. \implies y^2 + 2(x - c + d)y + x^2 - 2(c - d)x + 2cd - c^2 - d^2 = 0$$

This is a quadratic. Solving for $y_{1,2}$:

$$y_{1,2} = \frac{-2(x-c+d) \pm \sqrt{4(x-c+d)^2 - 4(x^2 - 2(c-d)x + 2cd - c^2 - d^2)}}{2} =$$

$$-(x - c + d) \pm \sqrt{x^2 + 2(d - c)x + c^2 - 2cd + d^2 - x^2 + 2(d - c)x - 2cd + c^2 + d^2}$$

$$y_{1,2} = -(x - c + d) \pm \sqrt{4(d - c)x - 4cd + 2c^2 + 2d^2}$$

Now we solve for when $x < c$:

$$x^2 + y^2 + 2cx - 2xy - 2dx - 2cd + 2cy + 2dx + 2dy - c^2 - d^2 = 0. \implies y^2 + 2(c + d - x)y + x^2 + 2(c + d)x - 2cd - c^2 - d^2 = 0$$

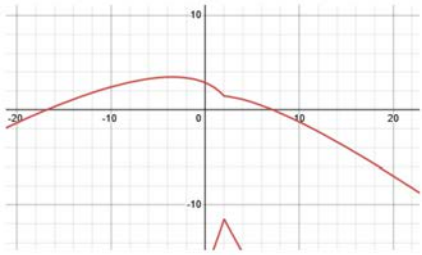
This is a quadratic. Solving for $y_{1,2}$:

$$y_{1,2} = \frac{-2(c+d-x) \pm \sqrt{4(c+d-x)^2 - 4(x^2 + 2(c+d)x - 2cd - c^2 - d^2)}}{2} =$$

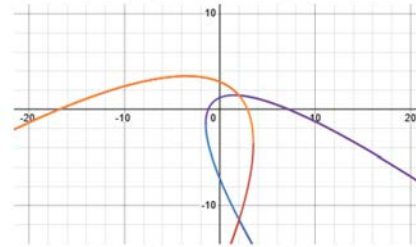
$$x - (c + d) \pm \sqrt{c^2 + 2cd + d^2 - 2(c + d)x + x^2 - x^2 - 2(c + d)x + 2cd + c^2 + d^2}$$

$$y_{1,2} = x - (c + d) \pm \sqrt{-4(c + d)x + 4cd + 2c^2 + 2d^2}$$

These solutions for y define parabolic branches that describe the boundary of the Dominance Region of Agent B $y \leq d$. It is therefore clear that this Dominance Region is unbounded. See Figure 2.6b.



(A) Dominance Region of Agent B with Agent A starting at $(2, 5)$ and $\gamma = \frac{1}{\sqrt{2}}$.



(B) Parabolic Branches Fit to the example in Figure 2.6a.

Figure 2.6
Example of Parabolic Branches fit to Dominance Region.

Each of the proposed four cases have been proven to be unbounded regions for Agent B, thus, Agent B's dominance region is unbounded for $\gamma \geq \frac{1}{\sqrt{2}}$ and Agent A starting at (c, d).□

Motion in Presence of Barrier

The effects of a finite straight-line barrier were studied in this exercise not to prove a theorem surrounding its effects, but to motivate further study through numerical examples. There are too many analytical and algebraic cases to make any broad conclusions about interactions of an agent whose movement is measured with respect to the L^1 metric at the moment. This is a first attempt to see how such a barrier affects an agent whose movement is measured with respect to the L^1 metric.

The Barrier's Definition and How It Affects the Agent's Isochrones

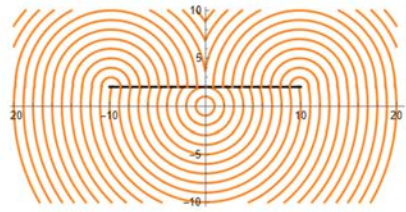
Introduce a barrier understood as a finite, straight line. To model this, the barrier is treated as an impassible region. Either target must move to the end points of the barriers and have its isochrones permeate out from there. The methods for this differ between the two means of describing movement.

DEFINITION 7 (Line of Sight). *The line of sight (LOS) is a linear curve of the form* $y = mx+l$ *that connects the observer at* (x_a, y_a) *and the observed at* (x_b, y_b).

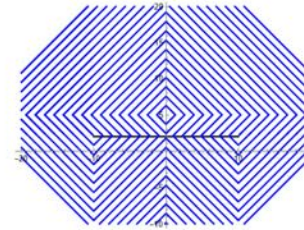
The L^2 case is discussed extensively in the paper Oyler et al. 2016, but it depends largely on the line of sight (LOS) of the agent to the edges of the barrier; a simple case is shown in Figure 3.1a.The region up to the lines of sight and barrier are treated the same as with no barrier. Past these lines, the isochrones are plotted as if the agent started from the edge of the barrier, with a difference in the starting time, t, to account for the time it takes to move to the edge of the barrier. The line where the isochrones intersect above the barrier (the y-axis above the barrier in this case) represents the set of points at which it would take the same time to reach traveling around the barrier either direction, when starting from the agent's initial location.

REMARK (Axis of Motion). *In the* L^1 *metric, the distance between points* (x_1, y_1) *and* (x_2, y_2) *can be decomposed as* $d_1((x_1, y_1), (x_2, y_2)) = d_1((x_1, y_1), (x_1, y_2)) + d_1((x_1, y_2), (x_2, y_2)) = d_1((x_1, y_1), (x_2, y_1)) + d_1((x_2, y_1), (x_2, y_2))$. *Hence, the movement of an agent described with*

respect to the L^1 metric may as well be decomposed into a horizontal and vertical component with their sums equating to the whole distance traveled. This observation gives rise to the notion of axes of motion: One axis of motion denoting a horizontal component and the other a vertical component.



(A) Simple case for Agent B's isochrones around a barrier.



(B) Simple case for Agent A's isochrones around a barrier (single AoM blocked).

Figure 3.1
The barrier's definition and how it affects the agent's isochrones

The lines along which the isochrones are affected is along one of these axes of motion. As in **Figure 3.1b**, the isochrones change along vertical lines on either side of the barrier as opposed to the line of sight with the center of the barrier as in the L^2 case.

Two terms will be used to describe the change that the barrier causes for each agent:

REMARK (Global Change). *Global change is defined as a change to the isochrones that does not correct itself with time, i.e, in **Figures 3.3a** and **3.3d**, the shape never returns to the original diamond shape.*

REMARK (Local Change). *Local change is defined as a change to the isochrones that is limited to a region directly behind the barrier (with relation to the agent). Such change is only possible for an agent whose movement is measured with respect to the L^1 norm with a barrier which does not block one of its axes of motion.*

Barrier-Isochrone Behaviors with Respect to the L^2 Norm

An agent whose movement is measured with respect to the L^2 metric will always have change to its isochrones with a barrier present, as the L^2 agent can no longer choose a straight line curve to a point past the barrier. Some cases are examined in **Figure 3.2**. Observe that there is global change to each of the isochrones past the barrier.

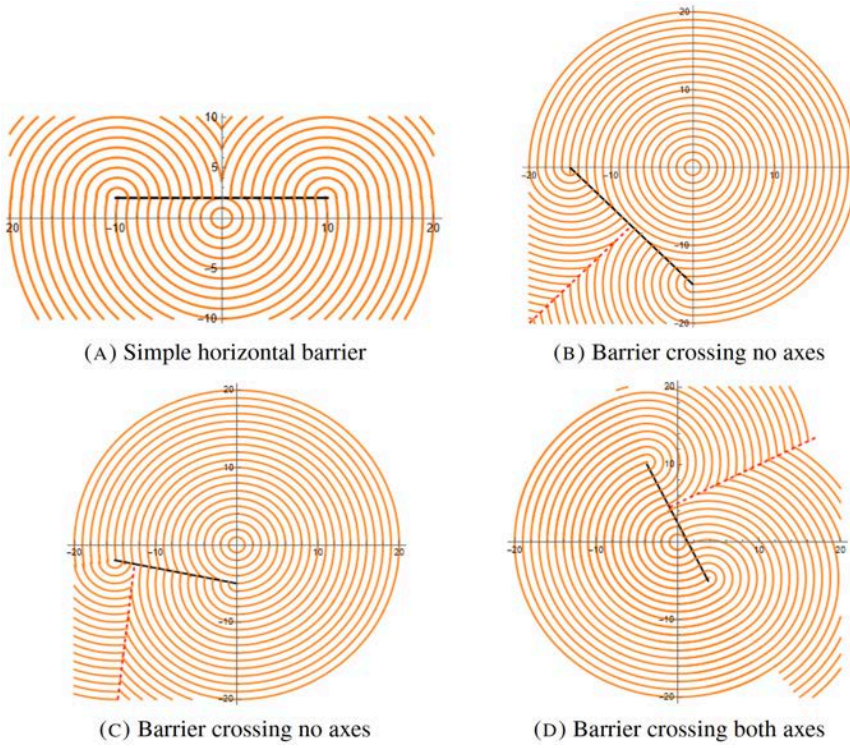


Figure 3.2
Various barrier positions and the effect on the isochrones of an L^2 agent.

Numerical Exploration for the L1 Metric

Effects of Barrier on Motion Measured with Respect to the L1 Metric

The addition of a finite, straight-line barrier affects an agent's, affects the isochrones and the resulting dominance region either locally or globally depending on its position with respect to the agent's axis of motion. This is represented in three cases: one axis of motion (AoM) blocked, both AoM blocked, and no AoM blocked. The case with single AoM and both AoM being blocked results in global change of the dominance regions, whereas the case with no AoM blocked only results in local change.

NOTE. We explore three separate cases: one AoM blocked, two AoM blocked, and no AoM blocked. These cases are examined for the effects on the isochrones and the resulting dominance regions. To start, the case with one axis of motion blocked:

Effects on the Isochrones

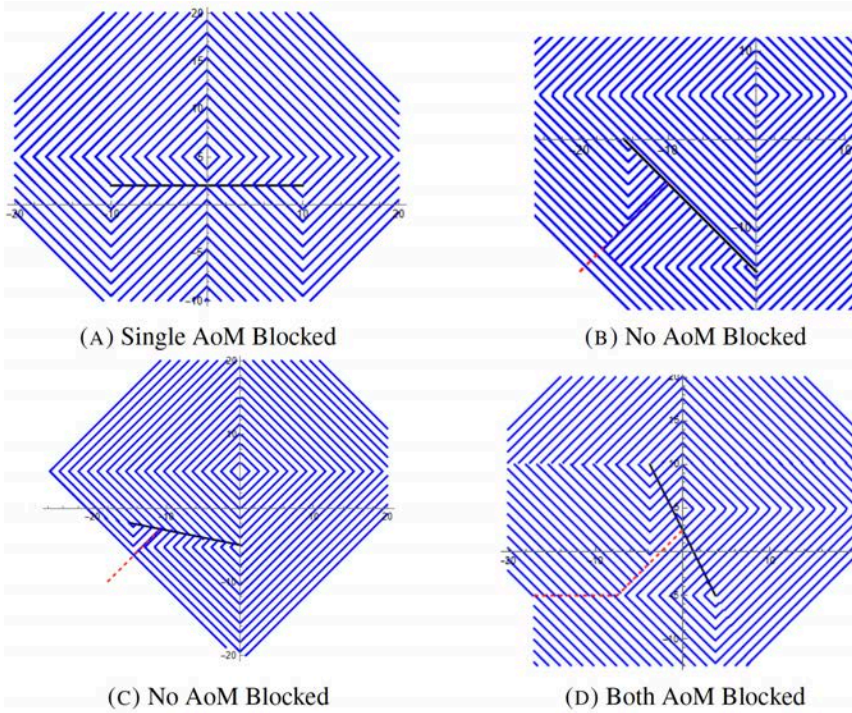


Figure 3.3
Various barrier positions and the effect on the isochrones of Agent A.

For **Figure 3.3a**, a single AoM is blocked, causing global change [3.1] along the negative y-axis. Its movement along the x-axis is unchanged, however. Observe that only the section defined by the vertical tangents on either side of the barrier is affected. The change to the agent's isochrones is still global because it will take longer go around the barrier and reach any point behind it, as it cannot move through the barrier. This global change is only present along the axis of motion that is blocked and is limited by the length of the barrier. Thus, there is global change to the agent's isochrones if one axis of motion is blocked.

Next, the effects of a barrier blocking no AoM is examined.

For **Figures 3.3b** and **3.3c**, there is local change [3.1] surrounding the barrier, but the change 'corrects' itself with time, i.e, the isochrones return to their original shape once the agent's traveled distance is greater than the vertical or horizontal legs of the barrier (if the barrier were treated as a hypotenuse). This is because Agent A may simply choose a path that goes completely around the barrier along either of its axes of motion. Thus, there is only local change for an agent's isochrones when the barrier blocks no AoM.

There is one more case, where both of the agent's axes of motion are obstructed.

In **Figure 3.3d**, there is global change along both axes of motion. This corresponds to change for both the vertical and horizontal components of the agent's movement past the barrier because it has to move around the barrier. Thus, there is global change along both AoM when the barrier blocks both of these AoM.

Effects on the Dominance Region

To determine how the dominance regions were affected by the barrier, the Dominance Regions were found in sections, with the sections where the agent is treated as starting from the edge of the barrier having a distance equal to that it would have to travel shifting it directly away from the other agent. This is because when parameterizing the equation, we removed time from the isochrone equations, thus we must treat the distance traveled (constant velocity times time) as the time difference in this case.

NOTE. All examples are graphed with a $\gamma 2 = 0.6$. The dominance region is affected in a very similar way to the Agent A's isochrones.

In **Figure 3.4b**, when compared to the base case in Figure 3.4a, there is visible global change in the dominance region only in the vertical direction of the barrier because it affects only that AoM. The change affects both agents: the region above the barrier is effectively deleted, adding to Agent A's dominance region. The region directly below the barrier is shifted downwards, corresponding to global change in the dominance region of both Agent A (Blue) and Agent B (Orange) as a consequence. Thus, there is global change to the dominance region along the axis of motion blocked by the barrier.

Next, the case when both AoM are blocked will be examined.

Figure 3.4c shows the case with both axes of motion blocked. When compared to the base case in Figure 3.4a, it is obvious that there is global change in both directions from the barrier. Agent A (Blue) benefits from the barrier slightly on the region just to the right of the barrier and the angled region below the barrier is shifted slightly up and to the right. In every other region, Agent B's (Orange) dominance region is expanded, however, enabling more options for paths used to escape contained within the unbounded region. Thus, there is global change to the dominance region along both axes of motion of Agent A.

Lastly, the case in which the barrier blocks no axes of motion is examined.

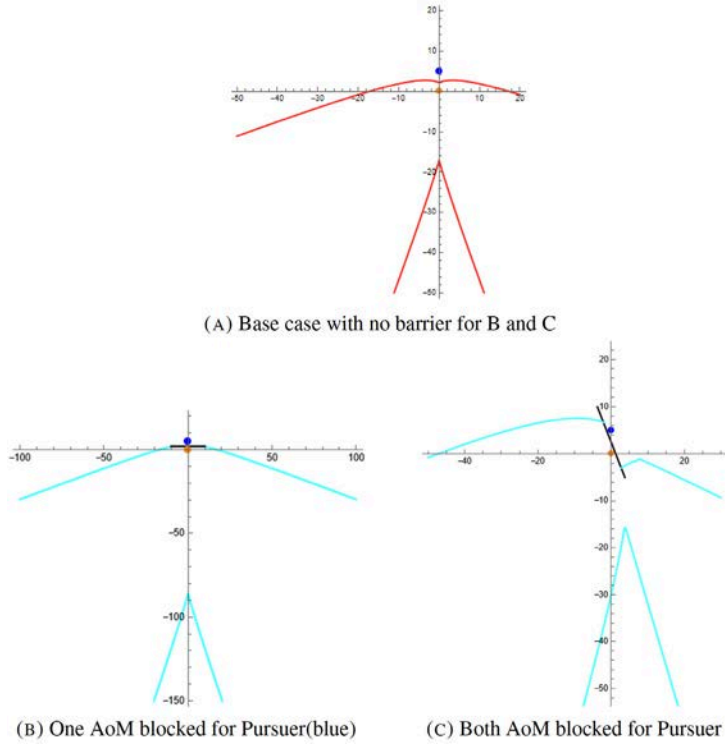
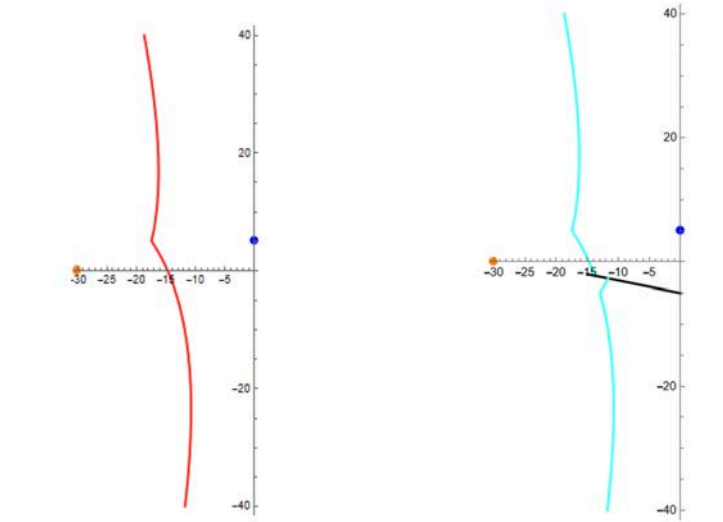


Figure 3.4
The dominance regions of a few of the cases from section 4.2

Figure 3.5b shows the case with no AoM blocked, and when compared against the base case (**Figure 3.5a**), it should be observed that there is only local change in the region directly behind the barrier with relation to Agent A. This corresponds to a slight increase in the dominance region of Agent B (Orange) and a resulting decrease in the region for Agent A (Blue). Thus, there is only local change in the dominance region when the barrier does not block Agent A's AoM.



(A) Base case with no Barrier for B (B) No AoM blocked for Pursuer

Figure 3.5
Base case and Barrier case for no AoM blocked

Conclusion

This paper entailed the exploration of escape and pursuit games, as well as the original work on limiting one agent to movement measured with respect to the L^1 metric and studying the effects of a barrier on such interactions. It was found that barriers affect an agent with movement measured with respect to the L^1 metric in a very different way than one with movement described with respect to the Euclidean (L^2) metric. The most significant finding was the L^1-L^2 Interaction Theorem, stating that for a velocity ratio of $\gamma = \frac{1}{\sqrt{2}} \approx 70.7\%$, a target is still able to escape if moving unrestricted (with respect to the L^2 metric). The second was a numerical exploration describing the barrier's effect on the isochrones and resulting dominance regions. This finding aligns with the first theorem. It provides insight into various applications and hints at a formal theorem.

Future Outlook

The current work lays the foundation for further work. Velocity was assumed to be constant and turn time/radius was not considered, but acceleration, both linear and angular, could be added in order to add complexity to the model and make it more applicable to realworld phenomena. The barrier is a finite, straight line, but the shape of the barrier could be generalized to a polygon or even a circle/ellipse. The movement could also be generalized to three dimensions, however this would require a lot of restructuring and would be rather difficult to visualize 3-dimen-

sional dominance regions and isochrones. Additionally, there is more work to be done to expand the exercise from this paper into a formal theorem, there are many possible algebraic possibilities for placing the barrier.

References

Nahin, Paul J. (1981). Chases and Escapes The Mathematics of Pursuit and Evasion. Princeton Puzzlers. Princeton University Press. ISBN: 978-0-691-15501-2.

Oyler, Dave W., Pierre T. Kabamba and Anouck R. Girard (2016). 'Pursuit–evasion games in the presence of obstacles'. In: Automatica 65, pp. 1–11. ISSN: 0005-1098. DOI: https://doi.org/10.1016/j.automatica.2015.11.018. URL: https://www. sciencedirect.com/science/article/pii/S0005109815004823.

Submission Guidelines

Eligibility

We consider papers on any research conducted by undergraduates at the U of M. Authors may be current U of M students, participants in U of M summer research programs, or students who have completed their undergraduate programs within the past three semesters. Papers may be single or co-authored, but at least one primary author must have been an undergraduate student at the time the research was conducted. In cases where students are members of a larger research team, the paper should reflect the students' contribution to the larger project.

Content

Submissions should not be creative manuscripts. Instead, they should present analysis of some type of data. In the case of collaborative research with faculty, papers should focus on the students' contribution to the larger project. While there is no minimum or maximum length for submissions, we understand that different disciplines have different norms, and that the length of submissions will vary. However, 8000 words or less is preferred. Any research involving human subjects should have received approval from the appropriate Institutional Review Board (IRB). Authors (in consultation with their faculty sponsors) are responsible for adhering to IRB guidelines.

Copyright

Submissions may not be under review for publication elsewhere. Authors retain copyright of their work and may publish their papers elsewhere after their Volume of *QuaesitUM* is published, provided the requirements of all parties are satisfied. Given the hard work that our editorial team puts into each paper prior to publication, if you publish your work elsewhere, we ask that you include an acknowledgement of *QuaesitUM* and its editorial staff.

Deadline for Submissions

Submissions are accepted anytime. All papers received by the first Friday of the spring semester will be considered for the volume being published in May/June. When possible, late submissions will be published the same year, but if not, they can be considered for the next volume.

Format

1. **Remove all occurrences of author(s) names** from the manuscript.

2. Submit a cover sheet **as a separate document** that contains the following information:

   (1) name of the author or authors,
   (2) date,
   (3) title,
   (4) citation style,
   (5) faculty sponsor*, and
   (6) your paper's category (according to the groups below)

   - Life and Health Sciences
   - Physical and Applied Sciences
   - Math/Computer Science/Engineering
   - Social and Behavioral Sciences
   - Humanities
   - Professional

   *The faculty sponsor is someone who supervised the research and provided guidance for the paper. They should agree to be contacted to recommend reviewers.

3. Upload an abstract of 150 words or less.

4. **Manuscript Formatting**

   - Upload your manuscript as a *.docx* or *.doc* file. (Let us know if your academic discipline uses a different format.)
   - Use size 12 Times New Roman.

- Include page numbers.
- Include a references page that follows the appropriate citation format for your field of study.
- Please single-space all submissions. Other than that, manuscripts should follow the formatting conventions for your discipline (italics, punctuation, capitalization, footnotes, citation style).

5. **Figures and Tables**

- Tables should be no more than 5 inches wide.
- Avoid extremely small letters/details whenever possible. These can become too small to read if we have to shrink the figure/table to fit within the 5"-page width.
- Provide a number and title for each figure and table.
- Place all figures and tables in your manuscript.
- Files must be a minimum of 300 pixels per inch (ppi).
- Also upload each table and figure in a separate file.
- Files must be a minimum of 300 pixels per inch (ppi).
- Figures should be in PNG, PDF, JPEG, TIFF, or EPS format.

6. Carefully check that each item in your references section follows the **punctuation, capitalization, and italicization** conventions of your citation style (APA, MLA, etc.).

Review Process

Submissions should be uploaded to the *QuaesitUM* submission manager. After submission, articles are screened by the *QuaesitUM* editorial staff. Manuscripts meeting and exceeding the panel's criteria are then forwarded to volunteer reviewers who have expertise in a field of research relevant to the manuscript. This is a double-blind process: both the authors and the faculty reviewers remain anonymous. Final selection decisions are left to the editorial staff.

Revision Process

After the review process is complete (usually mid-March), authors will be notified of the review board's decision. Authors whose papers are selected for publication will receive suggestions for revision from the faculty reviewers and editorial board. Authors are required to revise their manuscripts before final publication.

Award Opportunity

Exceptional papers will be eligible to receive a *QuaesitUM* Outstanding Paper Award, which carries a monetary prize of $500.

Submit Your Research Paper

Upload your manuscript and additional files at the *QuaesitUM* webpage:

memphis.edu/quaesitum

Questions and Further Clarification

Please contact us at **quaesitum@memphis.edu** with any questions.