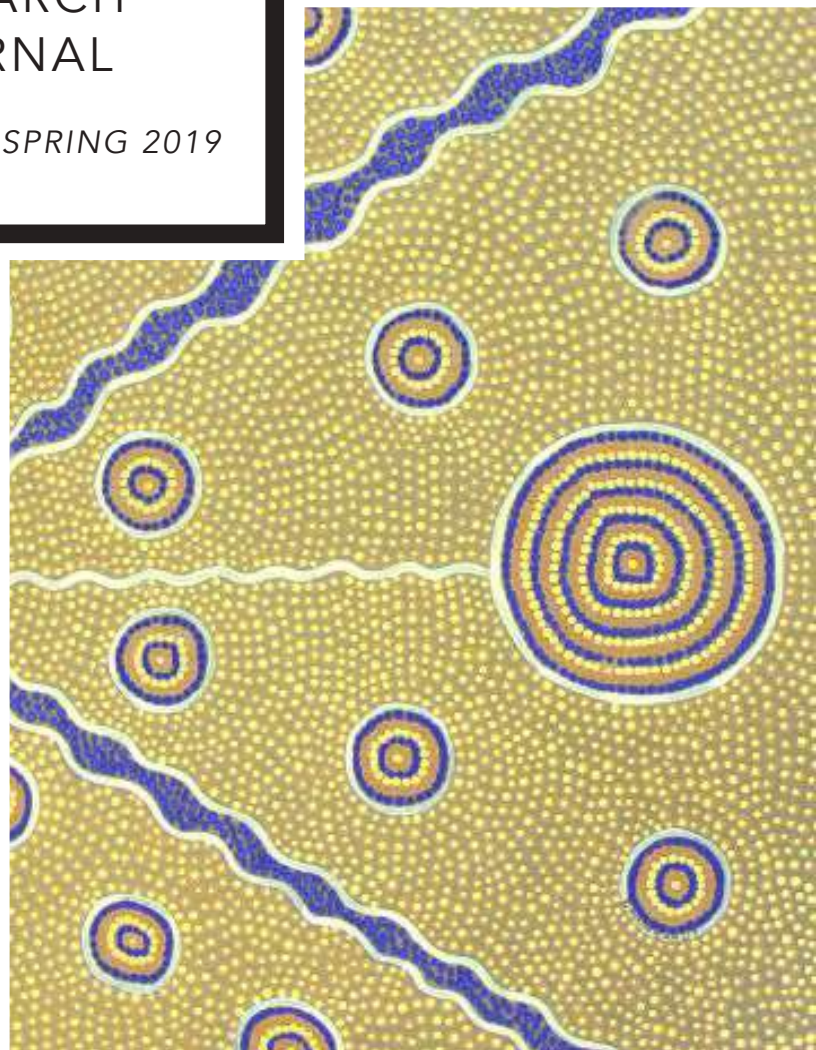
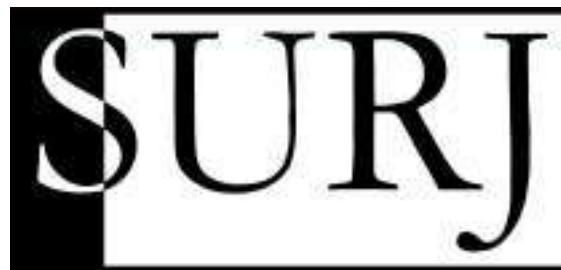


STANFORD
UNDERGRADUATE
RESEARCH
JOURNAL

VOLUME 19 | SPRING 2019





STANFORD UNDERGRADUATE
RESEARCH JOURNAL

VOLUME 19 | SPRING 2019

Stanford, CA 94305

INTRODUCTORY LETTER

Dear Readers,

We feel the need to open this letter with an acknowledgement of the role that critically understanding the world plays in understanding one another. Identifying past and present social structures that have defined our history and shape the modern day, as well as comprehending the natural world and our role within it, are crucial as we work to address the long-term issues that face us today.

In light of this, we are honored to present the eighteenth edition of the Stanford Undergraduate Research Journal. Since its founding in 2001, SURJ has produced intellectually vibrant collections of outstanding undergraduate research to celebrate students' accomplishments, share their discoveries, and promote an investigative spirit among undergraduates. We hope this new edition is yet another celebration of the pursuit of knowledge by students from Stanford and other universities around the world.

Understanding how to effectively communicate the discoveries made in the course of scientific research has become more critical than ever. In increasingly polarized and political environments, like policy discussions where research is seen as a tool exclusively to reinforce pre-existing beliefs, we recognize the importance of both science communication and open-access information in responding to the challenges the world faces.

As the repercussions of issues like climate change or voices that have been systematically ignored are increasingly acknowledged, we must ask ourselves how to discuss solutions in ways that are both logically and emotionally compelling. Research in this edition works to answer these questions. Many articles focus on identifying cycles of disempowerment and thinking about how voices and ideas that have been ignored or suppressed can be encouraged and strengthened. Other research analyzes patterns emergent from our collective behavior and how they can be better identified. A rigorous understanding of patterns in our decisions is a prerequisite to any action to improve people's lives, and one that the articles in this journal work tirelessly to satisfy.

This eighteenth volume of our journal is a large effort on the part of everyone involved. We would like to thank the ASSU, SAL, the undergraduate student body, and the research community on campus for your invaluable support for SURJ. And we would like to thank the SURJ editors' careful consideration and analysis as they worked with our featured authors to polish their works, our dedicated production team who painstakingly organized and structured these pages, and of course the authors who pursued answers to these fascinating and relevant questions.

On behalf of the 2018-2019 SURJ staff, thank you for reading this journal. We hope you enjoy the work inside.

Sincerely,
Eric Zelikman ('20) and Michelle Chang ('20),
Editors-in-Chief

STAFF DIRECTORY

"The mission of the Stanford Undergraduate Research Journal is to encourage, recognize, and reward intellectual activity beyond the classroom, while providing a forum for the exchange of research and ideas"

Executive

Editors-in-Chief

Michelle Chang
Eric Zelikman

Financial Officer

Varun Subramaniam

Production

Lead

Michael Kao

Associate Editors

Neeharika Bandlapalli
Woo Joo Kwon
JB Lim

Outreach

Lead

Eunice Yang

Outreach Associates

Betty He
Matthias Abebe
Roy Ghosh

Editorial

Social Science and Humanities Lead

Remi Akindele
Gulin Ustabas

Associate Editors

Matthias Abebe
Neeharika Bandlapalli, Betty He, Michael
Kao
Woo Joo Kwon
June Lee
JB Lim
Stephanie Shim
Jonah Reidel
Rizina Yadav

Natural Sciences and Engineering Lead

Sharon Huang
Manon Shah

Associate Editors

Tejas Athni
Alexis Froistad
Roy Ghosh
Rohan Mehrotra
Jonathan Sepulveda
Varun Subramaniam
Jenny Vo-Phamhi
Eunice Yang

Visit surj.stanford.edu to see past editions, how to join, and more

Microrhetorica: Ethos and Empiricism in Robert Hooke's *Micrographia* (1665)

PETER RAMSAY THOMAS
STANFORD UNIVERSITY

Robert Hooke's *Micrographia* (1665) offered the world a first glimpse into the microworld of various natural and artificial objects, from lice and moss to cork and written ink. Ostensibly, *Micrographia* championed the empiricism of the New Philosophy of seventeenth-century Europe that would mature into modern, natural science. Yet, beneath its veneer of objective, empirical inquiry, *Micrographia* reveals both the deep-seated social and moral anxieties of Restoration England and the exigent problems of credibility and authority that lay, and might still lie, at the heart of empirical science. I argue that Hooke's project in *Micrographia* was fundamentally rhetorical; its success and its very epistemic legitimacy were contingent as much, if not chiefly, upon a multifaceted ethos that strove to resolve paradoxes of personal credibility and impersonal objectivity that shaped the emergence of seventeenth-century natural philosophy. By studying Hooke's complex verbal and visual rhetoric, one discovers the achievements and compromises that enabled the heroic empiricism of the so-called "Scientific Revolution" and that might continue to govern modern assumptions about the nature and aims of scientific inquiry.

Introduction

Robert Hooke's *Micrographia* (1665) introduced to seventeenth-century England a minute world curious and previously unseen, revealed to its author by the recently-invented microscope and conveyed to its readers by thirty-eight engravings and extensive commentary.[1] The accuracy and detail of *Micrographia*'s depictions of plants, insects, and other objects, both natural and artificial, endure as testimony to Hooke's keen eye and illustrative skill, his surpassing capacity as an observer. However, the conceit of *Micrographia* is not exclusively empirical, but also, if not primarily, rhetorical. In order to present his empirical knowledge of microscopic bodies persuasively, Hooke employed both verbal and visual language to establish a bipartite ethos that supported both the credibility and the objectivity of his observations and himself as an observer. The first component of Hooke's ethos, resting on an associative frame of reference that drew upon the social and political prestige of Hooke's colleagues and patrons as well as the moral consensus of natural theology, secured his status as a trustworthy author. The second component drew upon the mimetic qualities of Hooke's illustrations, which he called Schemes, to diminish his authority over *Micrographia*'s contents and, therefore, to efface the impression of human artifice imposed on a stable, natural world created by the providence of an omniscient deity. Together, these opposing functions of Hooke's ethos served to explain a hitherto invisible microworld as, at once, the province of sophisticated empirical science and the flawless, orderly creation of God.

Text and Content: Establishing Social Credibility

The text of *Micrographia* fulfilled the first of these ethical functions, legitimating Hooke's claims to empirical knowledge by establishing his moral and social credibility. *Micrographia*'s twin dedications, the first to King Charles II and the second to the Royal Society

of London, served as initial and prominent vehicles for this legitimation. Although conventional, Hooke's regal dedication nonetheless laid crucial foundations for his credibility on political ground only recently (if then only precariously) settled following the tumult of civil war and republican dictatorship. Hooke insisted that the prosperity of "Philosophy and Experimental Learning," of which *Micrographia* was one of several fruits, was chief "[a]midst the many felicities that have accompani'd your Majesties happy Restauration and Government," and that foremost among the treatise's purposes was "to offer some of the least of all visible things, to that Mighty King, that has establish't an Empire over the best of all Invisible things of this World, the Minds of Men." [2] By presenting his work as both the product of monarchy and the means of its reinforcement, Hooke ensured that his discoveries would not at all subvert, but instead support, the project of stabilizing a restored but still tenuous society and its government. Such would be the only intention of gentleman, whose credibility would, thus, be nigh unimpeachable. Unlike his contemporaries, such as the chemist Robert Boyle, Hooke wanted for such gentlemanly credentials, which his birth to an Anglican minister failed to furnish.[3] Indeed, endorsing monarchy served to align Hooke with a social and political elite that could be trusted to offer reliable accounts of natural reality—especially important to Hooke who proposed to account for the invisible.

Micrographia's second dedication to the Royal Society of London, founded by Charles II in 1660 (the year of his Restoration) carried out a related rhetorical role. In confirming himself as a member of the Society and his present project as obliged to "those many Engagements [the Society] ... laid upon [him]," [4] Hooke once more arrogated to himself external epistemic authority, namely that housed within the Royal Society and, more precisely, its first Fellows. Despite the relative immaturity of the Society, and its consequent lack of institutional legitimacy, its early membership was certainly a

lack of institutional legitimacy, its early membership was certainly a “most illustrious assembly.”[5] For instance, not in the second dedication, but rather at the end of the subsequent preface, Hooke praised another of the Society’s founders, John Wilkins, as “a man born for the good of mankind, and for the honour of his Country... [whose] Zeal has been so constant and effectual in advancing all good and profitable Arts...” and who has, therefore, allowed “the chief Seat of generous Knowledge and true Philosophy” to reside in England.[6] He also acknowledged another colleague, the anatomist, astronomer, and architect Christopher Wren, as “so eminent a Person... who was the first that attempted any thing of this nature [i.e., to illustrate and annotate microscopical observations].”[7] And Hooke’s references to other members of the Royal Society spared neither in quantity nor extravagance beyond even the preface. In fact, *Micrographia* was peppered with praise for Sir John Cutler’s “munificence...in endowing a Lecture for the promotion of Mechanick Arts;”[8] for “the eminently Ingenious and Learned Physician, Doctor [George] Ent;”[9] for “the most Illustrious Mr. [Robert] Boyle...the Patron of Philosophy it self;”[10] and for the numerous other members of the Royal Society. Thus, the credibility and authority of the Royal Society, or at least of its esteemed Fellows, was endowed to Hooke by repeated affiliation. By their plentitude and their recursion, Hooke’s effusions serve not to align him and his work with a single gentleman philosopher, but with a cadre of learned squirearchs whose corporate methodology and association vis-à-vis the institution of the Royal Society constituted a veritable fountain of credibility from which Hooke was eager to drink. Hooke’s personal ethos in *Micrographia* was, thereby, thoroughly nourished, but so too was the collective ethos of the fledgling Royal Society. By portraying the moral unity of the Society’s membership, Hooke awarded to the body as a whole a persona of shared virtue that could authenticate the claims of each of its members, himself, of course, first among them. The basis for Hooke’s literary ethos as a trustworthy, gentlemanly observer, then, was associational. By associating his work and himself with credible authorities, he appropriated their integrity to himself. However, Hooke expanded his associative frame to encompass not only social and institutional sources of moral and epistemic authority, but also the religious and philosophical habits of mind that could afford him a similar legitimacy.

Natural Philosophy and Natural Theology

The detailed commentary on *Micrographia*’s illustrations was not disinterested, nor was it prosaic or purely descriptive; rather, it was moralizing and explicitly invested in the confirmation of a natural theology that explained a mechanical universe subject to a rational order imposed upon it by its Creator. Indeed, Hooke maintained that “so infinitely wise and provident do we find all the Dispensations in Nature...that [they] wil...appear the products of the highest Wisdom and Providence.”[11] The increasingly mechanistic natural philosophy of the seventeenth century was ever the close ally of natural theology, which promised a singular, functional order despite the restive tendencies of government, society, and organized religion.[12] In claiming that he, through *Micrographia*,

gave “infinitely cause further to admire the wisdom and providence of the Creator;”[13] Hooke cemented this alliance and extracted from it the moral authority necessary to validate his empirical observations. What is more, Hooke asserted that his microscopy prosecuted a particularly vital and virtuous task, revealing that “the Wisdom and Providence of the All-wise Creator, is not less shewn in these small despicable creatures, Flies and Moths...then in those greater and more remarkable animate bodies, Birds.” [14] Or that moss, which “the wisest of Kings thought...unworthy” is, in fact, “a most perfect Vegetable, wanting nothing of the perfections of the most conspicuous and vastest Vegetables of the world.”[15] The ethical value of Hooke’s theological and moral associations, as much as that of his social and political affiliations, is clear. They established Hooke’s credibility as a trustworthy, erudite gentleman whose *Micrographia* could legitimately claim knowledge about the minute bodies its author observed.

Crucially, by establishing a personal ethos, Hooke not only admitted, but also emphasized his role as *Micrographia*’s author. It was Robert Hooke who was credible, for his deference to monarchy; for his membership in the Royal Society; for his theological commitments; and for his moral rectitude. The authority of the observations contained within *Micrographia* could not, therefore, be divorced from Hooke’s authorship. His proximity, as an abstract, impersonal observer, to experimental phenomena was insufficient to impart legitimacy to purported empirical knowledge, for the recognition of that knowledge was also contingent on the “moral texture of social relations.”[16] Thus, Hooke exerted a magisterial, authorial presence through the text of *Micrographia*, offering incessant moral and religious commentary and frequent references to the achievements of his venerable colleagues in the Royal Society. Such a presence reveals Hooke’s, and, indeed, Restoration England’s, anxieties about moral and social credibility. But it also effectively resolves those anxieties with an associative rhetoric that substantiates Hooke’s authorial ethos to legitimate his figure as a reliable empiricist.

Visual Rhetoric and the Limits of Authority

But an authorial ethos alone was not adequate to the lofty epistemological status Hooke sought for *Micrographia*. Contingent as it was on the social and moral credibility Hooke endured himself, such an ethos threatened to conflate Hooke’s authorship of and authority over *Micrographia*, which he must have, with authorship of and authority over the microworld he had observed, which only God might have. What use was the frequent invocation of the “All-wise Creator” if control over creation was seized, rather, by Hooke? In order to maintain the integrity not of his persona but of Nature, Hooke established a second component of his ethos by which he became no author at all, but a personless observer. Already the text of *Micrographia* was dominated by Hooke’s authorial presence, necessary to validate his visual and verbal testimony. Thus, it is through *Micrographia*’s illustrations that Hooke worked to extract himself (or, his rhetorical self) from his own work.

To be sure, these illustrations were not isolated from

“
The detailed commentary on *Micrographia*’s illustrations was...moralizing and explicitly invested in the confirmation of a natural theology that explained a mechanical universe subject to a rational order imposed upon it by its Creator.”

the text; nor did Hooke draw rhetorical or conceptual boundaries between them. They are obviously complementary, not least for the indices Hooke used to label and organize the illustrations. Indeed, it was through his textual references to the illustrations themselves that Hooke accomplished part of his self-extractive task. For Hooke did not name *Micrographia's* graphics as illustrations, nor drawings, nor sketches, nor pictures, and certainly not as images, which would be, at best, faithful representations or, at worst, misleading illusions. He insisted, instead, on calling them Schemes, from the Latin *schema* or the Greek *σχῆμα* (*skhēma*) meaning “form.” [17] Hooke’s fidelity to the term *Scheme* might be read as mere terminological consistency. But it might also be interrogated as evidence of Hooke’s deep-seated anxiety about the artifice of images and as a sign of his efforts to dissimulate the identity of his schemata as such.

Although Hooke’s schemata remain images in fact, their presentation as forms accomplished two distinct, but related, goals, one practical and one ethical. Practically, as *Schemes*, Hooke’s illustrations acquired far greater epistemological power than they might as images. They become efficacious vehicles for what Steven

Shapin calls “the technology of virtual witnessing,” which “involves the production in a reader’s mind of such an image of an experimental scene as obviates the necessity of either its direct witness or its replication.” [18] The effectiveness of *Micrographia's* illustrations as media for virtual witnessing obtained, in large part, in the superior status Hooke awarded them as *Schemes*. The schemata remain mimetic; they are ultimately depictions, accurate as those depictions may be. Yet, by insisting on their formal integrity, Hooke separated the schemata from the agency of any artist (read: author), himself, who might distort or dilute them. Thus, Hooke’s *Schemes* served to erase, if not, at least, to obscure the role of Robert Hooke in their (re) production. The same Robert Hooke whose commanding, authorial ethos was omnipresent in *Micrographia's* text was, in *Micrographia's* *Schemes* no author at all, but merely an amanuensis copying the microscopic text of God’s creation in perfect simulacrum.

Ironically, however, Hooke’s schemata were hardly unaltered by his hand. While the microscopic details Hooke observed, described, and depicted were matters of anatomical and physiological fact, their presentation corresponded to an

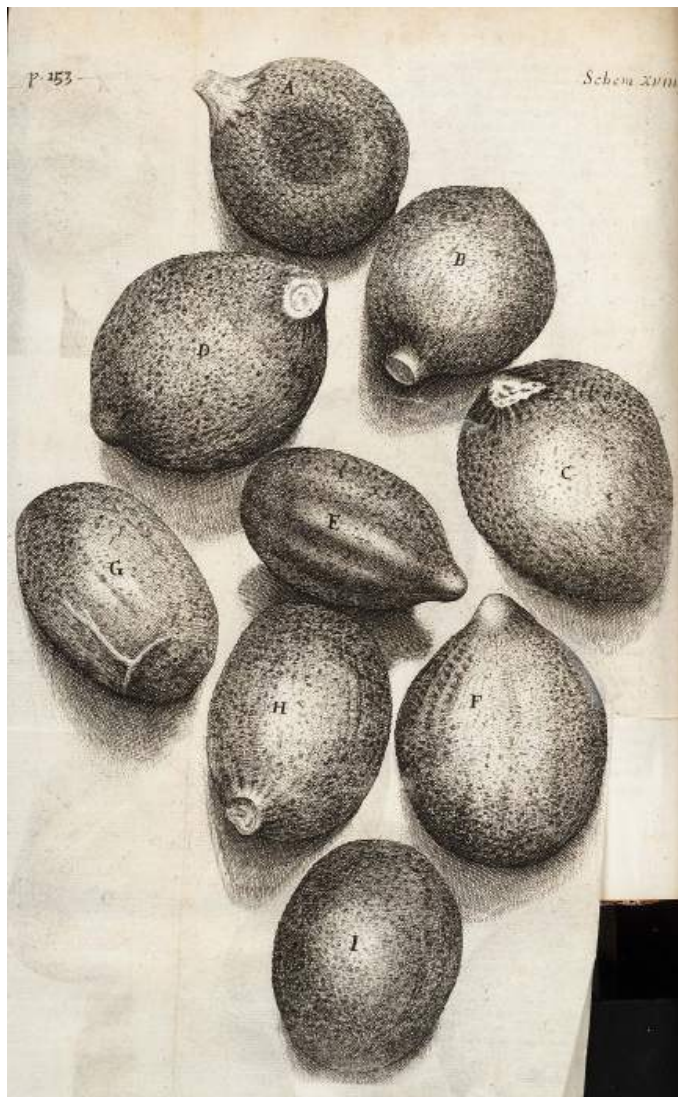


Figure 1. *Micrographia's* Scheme 13, a depiction of moss that expresses the formal characteristics of seventeenth-century naturalistic painting.



Figure 2. *Micrographia's* Scheme 18, a depiction of thyme seeds deliberately arranged and expertly rendered.

exaggerated visual rhetoric of order and stability informed by Hooke's mechanistic conception of nature and expressed by pictorial and compositional techniques borrowed from seventeenth-century artistic naturalism.[19] The engraving of moss (Scheme 13), for instance, betrays Hooke's debt to contemporaneous visual arts.[20] The special inclusion of a formal border frames the schema as a picture, while the extension of the moss fibers beyond and even over the frame accentuates their three-dimensionality, reproducing, albeit on a smaller scale, such a trompe l'oeil as might be expected on a Baroque canvas. Similarly, Scheme 18 captures thyme seeds in perfect stasis, their outlines briskly defined, their forms well-illuminated, and their shadows drawn in sharp contrast to highlight their lemon-like shapes.[21] Art historian Světlana Alpers notes the close similarities between such illustrations and the still-life paintings of the Dutch Golden Age. Indeed, Alpers's characterization of Hooke as the exemplar of a detail-oriented "descriptive impulse" in the seventeenth-century Northern European arts and sciences adequately explains Hooke's heuristic attachment to naturalism in his Schemes.[22]



Figure 3. *Micrographia's* Scheme 24, a depiction of a dronely displayed as calm and intact despite its prior dissection

But Hooke's reliance on artistic naturalism and the illusionistic techniques necessary to achieve it not only reflected a broader cultural agenda like that which Alpers and others describe, but also a rhetorical project that subtended the composition of *Micrographia*. In revealing a new, hitherto unseen level of creation (i.e., the microscopic), Hooke seized an opportunity to reinforce conceptions of the mechanical structure and fundamental order of the natural world. His uses of illusionistic perspective in his depiction of moss and of chiaroscuro (light-shadow modeling) in his illustration of thyme seeds, for instance, created an impression of a microworld organized according to the highest standards of realism that could be set for nature.[23] Further, by presenting his various flies and insects as intact, living organisms, rather than the mutilated and dismembered fragments they would have become following their dissection (procedures on which Hooke elaborates in great detail), Hooke strove to reconstruct microcosmic stability in

order to present the minute world as tranquil and unified. In Hooke's schemata, at least, the microworld achieves a discipline and calm long absent from the visible order of natural relations, be they anatomical and physiological or social and political. Moreover, by drawing once more upon established, external sources of philosophical and theological understanding, and reflecting in his Schemes a state of nature that those sources described, Hooke further retracted the shadow of his imperfect human hand from *Micrographia's* images, and therewith buoyed their status as special, visual capsules for microscopic components of the divinely-authored Book of Nature.[24] By presenting an orderly microcosmos, however constructed its order might be, Hooke presented a clear vision of nature consonant with the expectations of the seventeenth-century natural philosophy that would validate it.

Conclusion

The foremost achievement of *Micrographia* is, thus, rhetorical. Although his treatise contributed invaluable to anatomical, physiological, and microscopical knowledge, Hooke's simultaneous self-presentation as the illustrious, involved author of his book and as the humble, marginal amanuensis of God's creation (i.e., nature) established enduring conventions for the exhibition of that knowledge. Hooke's portrait of the Royal Society, description of its methodology, and publicization of its philosophic exploits, not least among the recognized accomplishments of *Micrographia*, hinged as much on Hooke's rhetoric as on the coherence and utility of the knowledge he gathered.[25] On the one hand self-emphasizing and, on the other, self-effacing, Hooke's composite ethos ensured that *Micrographia* would be received as a true account of a theretofore unseen stratum of creation, both for Hooke's integrity as an experimental philosopher and for his personal distance from the artifices of experimentation. Rather than self-defeating and contradictory, Hooke's ethical synthesis was self-reinforcing and robust. To read *Micrographia* as a rhetorical artifact is to reveal the interplay between credibility and objectivity in the presentation and, even, construction of empirical knowledge—an interplay that lies at the heart of the methodologies and sensibilities of modern science. For, ultimately, Hooke's dependence on rhetoric did not so much cripple his scientific project as it did constitute its very essence.

References

- [1] Robert Hooke, *Micrographia: or Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses. With Observations and Inquiries Thereupon*, facsim. ed.. (1665; repr., Lincolnwood, Ill: Science Heritage, 1987).
- [2] Hooke, *Micrographia*, ii-iii.
- [3] Steven Shapin, "Who was Robert Hooke?," in Robert Hooke: New Studies, eds. Michael Hunter and Simon Schaffer (Woodbridge, Suffolk: Boydell Press, 1989), 256.
- [4] Hooke, *Micrographia*, iv.
- [5] Ibid.
- [6] Ibid., xxxii.
- [7] Ibid.
- [8] Ibid., xxxi.
- [9] Ibid., 105.
- [10] Ibid., xviii.
- [11] Ibid., 177.
- [12] See Michael Hunter and Simon Schaffer, "Introduction," in

Robert Hooke: New Studies, eds. Hunter and Schaffer, 1-19.

[13] Hooke, *Micrographia*, 193.

[14] *Ibid.*, 198.

[15] *Ibid.*

[16] Steven Shapin, *A Social History of Truth: Civility and Science in Seventeenth-Century England*, Science and Its Conceptual Foundations (Chicago: The University of Chicago Press, 1995), 381.

[17] Oxford English Dictionary, 2nd ed., s.v. "scheme" (Oxford: Oxford University Press, 1989).

[18] Steven Shapin, "Pump and Circumstance: Robert Boyle's Literary Technology," *Social Studies of Science* 14, no. 4 (1984): 491.

[19] Hooke's interest in naturalism and basic familiarity with it may be explained by his apprenticeship to royal portraitist Peter Lely and to his life-long appreciation of the pictorial arts. See Lisa Jardine, *The Curious Life of Robert Hooke: The Man Who Measured London* (London: HarperCollins, 2003); Janice Neri, "Between Observation and Image: Representations of Insects in Robert Hooke's *Micrographia*," *Studies in the History of Art* 69 (2008): 82-107.

[20] Hooke, *Micrographia*, 130a.

[21] *Ibid.*, 160b-c.

[22] Světlana Alpers, *The Art of Describing: Dutch Art in the Seventeenth Century* (Chicago: The University of Chicago Press, 1983), 84-86.

[23] See Eileen Reeves, "The New Sciences and the Visual Arts," in *A Companion to Renaissance and Baroque Art*, eds. Babette Bohn and James M. Saslow (Oxford: Wiley Blackwell, 2012), 316-335.

[24] The natural theology upon which Hooke drew was widespread and popular especially in England. Its exemplars, many of them Hooke's peers in the Royal Society, included John Wray (*The Wisdom of God Manifested in the Works of the Creation*, 1691) and William Derham (*Physico-theology, or a Demonstration of the Being and Attributes of God*, 1713).

[25] On Hooke's later influence, see John T. Harwood, "Rhetoric and Graphics in *Micrographia*," in *Robert Hooke: New Studies*, eds. Hunter and Schaffer, 119-148; Jardine, *The Curious Life of Robert Hooke: The Man Who Mapped London*; Shapin, "Who was Robert Hooke?"

Representational Abstractionism and Realism: Two Approaches to Chicano Art and Literature

MIGUEL SAMANO
STANFORD UNIVERSITY

Chicano art has been described as a canon of representation distinct from mainstream American art. After all, Chicano artists often represent different subjects than mainstream artists, especially those artists who were active during the Chicano civil rights movement. But how distinct are their strategies of representation from those of the mainstream? Tomás Ybarra-Frausto, among others, takes a representationally realist approach in presuming that one of the distinguishing features of Chicano art is the verisimilar representation of the experiences of Chicanos. In this paper I argue that Chicano art can also be described as critiquing the contingent nature of social relations, such as holding a collective identity as Chicano, by drawing attention to the necessarily arbitrary and distinct nature of representation, just as mainstream abstract art does. I term this position representational abstractionism. First, I argue through a reading of Ybarra-Frausto's essay "Rasquache: A Chicano Aesthetic Sensibility" that his nominally descriptive account in fact prescribes a representationally realist aesthetic to Chicano art. Then, I develop a reading of Raúl Salinas' poem "Un Trip through the Mind Jail" that demonstrates how the search for a collective identity as a Chicano would be better thought along representationally abstractionist lines. This framing of Chicano identity as a contingent product of representation, I conclude, enables a different aesthetic politics, namely it renders the politics of authenticity incoherent.

The Chicano Movement of the late 1960s and early 1970s made a new form of collective identification available to peoples of Mexican descent living in the United States. In the domains of visual arts and literature, Chicanos set about transposing the sociopolitical project of the Movement into the realm of aesthetics, mainly through replacing external depictions of Chicanos within Anglo-American culture with representations of their own making ("The Chicano Movement" 129). This aesthetic project was accompanied by the formation of Chicanoist literary and artistic criticism that attempted to interpret, describe and historicize this project in terms of its aesthetic strategies. Early critics made Chicano art and literature recognizable as distinct canons of representation through descriptions and interpretations that reached beyond disinterested analysis, leading to the validation and dissemination of these works throughout the larger world of art criticism and the academic reading public.

In their articulations of an aesthetic project, critics constituted Chicano art as its own representational space by insisting on its distinctness from mainstream American art. The art historian and literary critic Tomás Ybarra-Frausto, for example, opposed the Chicano art movement to mainstream art movements because the former goes beyond "the level of the work" to "extend meaning beyond the aesthetic object to include transformation of the material environment as well as of consciousness" ("The Chicano Movement" 141). In context, Ybarra-Frausto writes from the standpoint of an art historian, which is to say his explicit aim is to describe an aesthetic, not prescribe one.¹ However, one may then ask if the division he

proposes between art that "extend[s] meaning" and "art about itself and for itself," is not, in fact, prescriptive in its insistence on the distinctness of Chicano art (141). For through its dwelling on its own constructedness, mainstream abstract art also often calls into question the constructedness, contingency, and therefore changeability of social realities.² Ybarra-Frausto's descriptive claim, then, works in part through prescribing a presupposition about art in general, that it usually does not engage the social world.

The formation of a representational space for Chicano art and literature occurred, on the one hand, through the prescriptive force of a descriptive aesthetics that insisted on the difference

alternative artistic discourses. For 'Hispanic' art, this selective incorporation often foregrounds artwork deemed 'colorful,' 'folkloric,' 'decorative,' and untainted with political content" (146). His attempt to provide an account of Chicano art with "consistent and defining stylistic features" (146), especially when considered contrapuntally with the mainstream power structure he describes, both describes and, in my view, prescribes an aesthetic, but through a different mechanism than the one I lay out in this essay. This mechanism appeals to the place of Chicano art and art history within an aesthetic field, bestowing upon it the valence of institutional critique. The positioning of Chicano art vis a vis mainstream art is outside the scope of this essay. Within the aesthetic field of the American artworld in the latter half of the 20th century, it is nonetheless true that movement-based ethnic art, such as Chicano art was often derided as "illustrative," "essentialist," and "cultural nationalist" by mainstream art critics (García; cited in Noriega 21), as art that "never rose to the challenge of modernism and its investment in aesthetic autonomy, formalism, individualism, and internationalism" (21). If Ybarra-Frausto's comment could be interpreted as positioning Chicano art as a postmodern practice that rejects "art about itself and for itself," the contrast he is drawing may function as a critique of historiographical perspectives that locate Chicano art outside of American art history as opposed to alongside its postmodern contemporaries. This is not incompatible with a reading of his comment as also drawing a contrast against abstract art, especially since the art practices he goes on to describe in his essay are predominantly figurative, not abstract, engaged, as Bourdieu might put it, with a primary referent outside of the universal of stylistic possibilities that is the history of artistic form, i.e., the comparatively autonomous realm of art history.

1 In "The Chicano Movement, the Movement of Chicano Art", Ybarra-Frausto implicitly responds to the exhibition *Hispanic Art in the United States*, which caused controversy within the Chicano community. "[T]he power structure of mainstream art journals, critics, galleries, and museums," he writes, "selectively chooses and validates what it projects, desires, and museums as constituent elements of various

between Chicano art and literature, and, on the other, mainstream art and literature. This was achieved through assumptions made in descriptive analyses, which were not about Chicano art per se but about the nature of art itself. This paper aims to question one of these assumptions, that the project of Chicano art proceeded on representationally realist terms, that what artists sought to accomplish through their work was the more or less verisimilar representation of Chicano lived realities. It puts pressure on this assumption through a close reading of Ybarra-Frausto's catalogue essay, "Rasquachismo: A Chicano Sensibility" (1991). The goal is not to show that all Chicano artworks are representationally abstractionist, but to point out how description may condition the readings of certain Chicano artworks as realist even when they are not. After my reading of "Rasquachismo," I will offer a representationally abstractionist alternative to Ybarra-Frausto's realist reading of "Un Trip to the Mind Jail" by the Chicano poet Raúl Salinas.

By *representational abstractionism* I do not mean the tendency in Western art since the late nineteenth century towards aesthetic abstraction as opposed to figuration, but the recognition that artworks are abstract by definition in their necessary distance from the social reality they are understood to represent. Artworks achieve an abstractionist effect by calling attention to their own artificial character in a move that invites us to question not only the aesthetic representations proffered in art but the social facts to which they allude.³ Art is abstract, but it's also embedded in the world. For example, Marcel Duchamp's *Fountain* (1917) is abstractionist in the sense that its readymade form impelled viewers to interrogate their assumptions about the nature of an artist's creative labor. The artworks of the Impressionists would also count as abstractionist under my definition since they drew attention to the prevailing theory of art by not aiming at the imitation of real life pointing up its contingency to expand the meaning of art to include works whose artistry lay precisely in their non-imitation of life (Danto 573-75). Because Impressionism focused on perceptual experience over and against the mimetic representation of a reality prior to it and recognized by an art audience, it was attuned to present moment and change rather than oriented towards the past. Impressionist art was as self-consciously constructed as Duchamp's ready-mades.

One may, at this point, readily object that this was precisely the distinction Ybarra-Frausto was aiming to draw between mainstream art and Chicano art. Impressionist art and ready-mades are self-consciously constructed, but the social facts they call into question pertain primarily to the art world. In contrast, Chicano art engages with social facts that go beyond it to address issues of broader political, social, and cultural concern. The case for Chicano art, as with other forms of ethnic American art, must be different. The relevant sense in which it is different seems to be that Chicano art as a category presupposes a correspondence between art and the set of social facts pertaining primarily to Chicanos such as biculturalism, barrio culture, urban youth culture, and so on. Abstractionism in Chicano art in this sense is referential in a way that mainstream art is not always and that accounts for descriptive aestheticians, such as Ybarra Frausto, reading aesthetic strategies as realist instead of abstractionist. Philip Harper phrased the conundrum for an African-American abstractionist aesthetic as follows, though his observation

holds for Chicanos as well: "how can a work clearly enough ground itself in the real-world racial order as to register as *black* while at the same time clearly enough *dissociating* itself from lived reality as to register as productively *abstractionist*." As Harper argues, the referential nature of African-American abstractionism makes that "lived reality" "available for critical interrogation" but only if instead of embracing realist norms for the interpretation of art such as an assumed verisimilitude, we embrace a set of abstractionist protocols. This often involves a sustained scrutiny of a text's claim to realism, as I will demonstrate through my reading of Ybarra-Frausto's catalogue essay on rasquachismo.

Rasquachismo as an aesthetic sensibility is realist in its insistence on tying Chicano urban working-class ("barrio") material culture and practices to the aesthetic strategies deployed by Chicano artists. Its claim to realism is premised on its comprehensiveness since Ybarra-Frausto posits it as a "vernacular idiom encoding a comprehensive worldview" "rooted in Chicano structures of thinking, feeling and aesthetic choice" ("Rasquachismo" 155). As such rasquachismo in art is for him aestheticizing a set of practices observable in everyday life in the barrio. He underlines this view rhetorically through juxtaposing descriptions of these practices outside of art with their particular adaptations in a variety of art media including drama, installation art, and narrative fiction. At the most macroscopic level, he achieves this juxtaposition by transitioning in the first section of the essay from a general overview of rasquachismo ("very generally, rasquachismo is an underdog perspective...") to identifying features of it in everyday practice in a manner that he explicitly flags as associative and not exclusively about art: "what follows, then, is a nonlinear, exploratory, and unsolemn attempt to track this irrepressible spirit manifested in the art and life of the Chicano community" (156). The following and final section of the essay, titled "historical continuity," then provides a series of examples across genres and periods of art that exemplify the rasquache sensibility (156-160). In terms of the essay's general structure, Ybarra-Frausto makes the relatively uncontroversial point that artists who see themselves as part of the same tradition by virtue of membership in the same culture often deploy similar strategies and that these strategies help their various artworks cohere together as a tradition. However, at the paragraph-level in the first section of the essay, his juxtapositions of everyday practices and their consequences for art imply a more determinative role for the former: Art does not so much call the rasquache sensibility practiced in everyday life up for comment as exemplify it. Ybarra-Frausto writes, for example, that part of the rasquache sensibility are a set of coping strategies ("movidas") that connote "resilience and resourcefulness" within the barrio; these strategies ("the use of available resources") engender the aesthetic strategies of "hybridization, juxtaposition, and integration" (156). Subsequent paragraphs in the first section follow a similar pattern of commenting on an aspect of a rasquache sensibility before drawing a conclusion about the strategies it engenders in art. In at least two instances this realist patterning leads Ybarra-Frausto to overlook objections to a cross-generic aesthetic sensibility based on the differences between particular genres.

The first instance occurs when he describes the visual distinctiveness of the barrio as projecting a "visual biculturalism" through its pairing of traditional items from Mexican popular culture with those from American mass culture (157). This biculturalism and the "visual interplay" it seems to effect is characteristic of some artworks in the *CARA* show, which primarily focused on visual artwork, because the spatiality of visual art readily facilitates this

³ Social facts are contingently true by a virtue of a social group taking them as such. An example of a social fact may be the statement "there are two genders," which evaluates as false among some social groups. In contrast, a fact such as "the sun rises every morning" may be contingently true, but its truth value is independent of any given social group evaluation of it.

“interplay” (157). In contrast, the “interplay” would be harder to achieve in linguistic genres due to the linear progression of language through time. As Saussure explains, “auditory signifiers [such as those of language] have at their command only the dimension of time. Their elements are presented in succession; they form a chain” which is why language attempting to describe visual phenomena can only do so incompletely (70). Linguistic genres pose a challenge to *rasquache* as a realist aesthetic insofar that attempts to render a biculturalism indexed specifically to the *barrio* will be limited by their need to excerpt specific scenes from *barrio* life, or, even, parts of scenes. The logic of language calls attention to the contingent and therefore artificial nature of excerption from a *barrio* lifestyle that is experienced in both space and time. Language as such remarks the distance between, on the one hand, the representation of the *barrio* proffered by Ybarra-Frausto and, on the other, the *barrio*’s reality. Yet we can readily imagine that an easy objection to this argument would be that language in its everyday use often does aim to represent the non-linguistic world as accurately as possible. Certain linguistic artforms, however, draw attention to their distance from everyday language use with one in particular often being characterized by that feature: poetry.⁴

The second instance when Ybarra-Frausto overlooks generic differences between art media brings to mind the formal logic of poetry. In his discussion he addresses the formal elements of altar composition — “precise repetitions, replications, and oppositional orders of colors, patterns, and designs”— to demonstrate the *rasquache* preference for composite organization (157). However, he does not address how it is that the individual elements of an altar cohere together to form the meaning of the composition as a whole. He does, though, provide a list of possible altar elements whose organizing principle seems to be metonymy: “plaster saints, plastic flowers, bric-a-brac, family photographs, and treasured talismans” (157).⁵ Each element contributes to the religiosity and intimacy of the home altar through its standing in for either an aspect of religious devotion or a memory in a family member’s life. Notwithstanding that each element of a composition may be loosely associated with many other elements outside of it, the composition’s meaning as a whole—as a home altar—depends on the ability of each of its elements to be understood as securing a single referent. The referential logic of metonymy as a trope supersedes associative, metaphorical interpretations of the altar. Yet to be convinced that this is always so for Chicano compositions of any genre, one would have to ignore that metaphor often is the trope that distinguishes poetry from other artforms (Jakobson). Though metonymy serves a similar function of securing stable referents for poems—and this is evidently the case for socially-engaged poetry such as Chicano poetry—poems are distinguished from other types of writing by

4 The sensuality of poetic language could be the spatial dimension of poetry. Language can be spatial, or textual, and not just temporal in this sense, but our treatment of language may nonetheless mute its sensuality.

5 Metonymy could also be applied to the analysis of visual art. An extremely metonymic composition, for example, may have its features index particular objects or events outside of the artwork. Visual metonymy seems different than linguistic metonymy, though, despite, in both cases, metonymy and metaphor being thinkable as two poles for representation. One difference may be that visual signs, whether symbols, indices, or icons, tend to impel viewers to locate the referent of the sign in a way that linguistic signs do not as readily. In my view, this is why some audiences find abstract art more frustrating than experimental writing, even when both no longer aim at representation in the sense of there being a readily recognizable referent to locate. If I am correct about the distinction between the visual and the linguistic with regards to metonymy, then Ybarra-Frausto’s claim that *rasquachismo* is a cross-generic aesthetic sensibility, as opposed to primarily a visual one, is a rhetorical overreach.

their relative open-endedness of meaning. If metaphorical language more generally works by casting one object in terms of another, metaphors in poetry account for the propensity of poems to open up new perspectives on reality. The logic of metaphor in poetry, then, can urge readers to denaturalize seemingly necessary relationships between particular representations and reality as ones among many—as arbitrary.⁶ In contradistinction with the example of a *rasquache* composite organization Ybarra-Frausto provides, poetry insists that any relationship between individual elements, the composite whole, and the social world may be secured as part of a *rasquache* aesthetic sensibility only if one underplays the relationship’s arbitrariness.

The distant and arbitrary relationship between representation and reality that poetry discloses suggests that the possibility of *rasquachismo*, as Ybarra-Frausto describes it, depends on presupposing representational realism. One must already be committed to believing that the aesthetic representations of Chicano art more or less correspond to the set of social facts about the *barrio* and its people. Ybarra-Frausto takes this presupposition as justified by the history of Chicano art since the Chicano Movement:

“Turning inward to explore, decipher, and interpret elements from the Chicano cultural matrix, artists and intellectuals [during the Chicano Movement] found strength and recovered meaning in the layers of everyday life practices. The very essence of a bicultural, lived reality was scorned as un-American by the dominant culture... a necessary response was to disown imposed categories of culture and identity and to create a Chicano self-vision of wholeness and completion” (159).

The representationally realist commitment lies in the response of displacing “imposed categories” with a “Chicano self-vision” drawing on a shared cultural matrix, evacuating “culture and identity” of their previous contents as opposed to calling the coherence of either category into question. Indeed, that the self-vision is one of “wholeness and completion” as opposed to fragmentation and indeterminacy suggests that the project of Chicano art must be capable of delivering up a coherent set of representations.⁷

6 Metaphors can also imply natural relationships between objects. If I were to describe Langston Hughes’ poems as “bluesy”, for example, I am implying a relationship between his poetry and the blues that could be construed as natural. It is not entirely clear to me if “bluesy” naturalizes the relationship on account of being metaphorical — in which case, the new perspectives that metaphorical descriptions disclose may seem astoundingly obvious in hindsight — or of being a particular kind of metaphorical description. I suspect that the latter account better explains how terms such as “bluesy”, “funky”, “*rasquache*”, “folkish”, and “urban” work. These terms make the material basis of linguistic representation (i.e., that there is a referent) clear because there is an implicit over-determination of what the (“natural”) referent of the metaphor is, i.e., Black and Chicano culture. For a discussion of Black music as indexing Black experiences, see chapter 2 of Phil Harper’s *Abstractionist Aesthetics* (NYU, 2015). I am not interested in reading racialized metaphors as pointing us to the same, tired referents. Rather, I want to reclaim the force of these metaphors by pointing out how these metaphors can also de-naturalize relationships if adopt an abstractionist protocol that insists on recognizing their arbitrariness (not wholly arbitrary). This will become clearer in my discussion of the poem “Un Trip through the Mind Jail”.

7 There is a difference between offering new representations within the same categories and challenging the concept of categories as such. The latter strategy, which was common among early Chicano conceptual artists, is arguably not so much an *identity politics* as an interrogation of the context for thinking about identity that the categories make possible. As Chon Noriega explains, the conceptualist art collective *Asco*, which was active and prominent during the Chicano Movement employed this strategy: “For *Asco*, what made Chicano identity performative was not that it named itself against all odds, as an act of defiance (the oft-used “I am Chicano” of the era), but that it was constituted within a set of social relations largely defined by the mass media and the corporate liberal state. As such, *Asco* saw identity as less a question of form and content — that is, a proper name, and more one about the *context for speaking and being heard*.” (24; my emphasis).

The formation of a Chicano self-vision achieved through a consistent refusal to accept social facts as necessary truths—including those structuring a sense of self and community—is the abstractionist alternative Ybarra-Frausto overlooks. In individual artworks, this may manifest in a divergence in the meaning contributed by cultural content and its embedding in particular forms wedded to particular logics. That at least Raúl Salinas employed the above strategy in his poetry suggests that the collective project Ybarra-Frausto aimed to describe may, in fact, have admitted more of abstractionism than he presumes. Moreover, in “Un Trip through the Mind Jail” Raúl Salinas enacts this strategy precisely through the conscious failure of the particular representations of barrio culture he invokes to overcome the distant and arbitrary nature of aesthetic representations as such. Thus, a reading of “Un Trip” oriented towards recognizing its abstractionist character casts doubt on both Ybarra-Frausto’s reading of it as “an attempt of showing how [Raúl Salinas’s] experience might be typical and representative of the Chicano community” and his presupposition that representations of barrio culture generally are realist (“Introduction” 10). The interpretive question motivating an abstractionist reading of the poem is whether or not the speaker accepts the limits that linguistic representation poses for his attempts to represent the past.

The poem progresses towards its end as a series of vignettes of life in a neighborhood that the poetic voice assures us was his own during his youth. Presented through a series of stanzas all related to each other through their refrain of “neighborhood of,” the speaker’s reconstructed past dominates the theme of the poem even as the necessary serio-temporal progression of poetry points out how this past is lost in time. The images on offer, though, are not simply temporal elaborations but spatial ones since through his invocation of particular characters in delineated situations he seeks to create the sense of a place, namely his neighborhood of La Loma, Austin. Among many other things, he tells us, La Loma is peopled by “[m] odest Mexican maidens dancing,” “kids barefoot/snotty-nosed,” and “girls from cleaner neighborhoods” who can be seen at, respectively, a “dilapidated community hall,” on “muddied streets,” and at parties at “Guadalupe Church” (55-56). Yet the incessant march of the poem forward leaves each vignette-stanza to flicker out, leaving the neighborhood of La Loma as only a loose series of associations for the reader, and its people as, at best, only partially reconstructed subjectivities.

As the poem draws to its end, the speaker seems to reconcile himself to this necessary incomplete recollection of the past borne out of its distance from the present. If the associative recollections of the La Loma he once knew cannot recuperate the neighborhood’s existence, he affirms that the act of recollection in the present is nonetheless all too real. He does so while explicitly invoking the on-going poetic present in the most self-conscious moment of the poem. He remarks that “only the NOW of THIS journey is REAL!”, where the capitalization of “NOW” and “THIS” draw attention to their grammatical function of marking a moment of enunciation which enables identity (re)formation through poetry (59). The

poem continues to remark the difference between the speaker as an enunciating subject and a subject of one of his recollections, when the speaker directly addresses the neighborhood: “neighborhood that is no more/ YOU ARE TORN PIECES OF MY FLESH!!!!” (59). The subject alluded to by the possessive pronoun “MY” is embodied. His persistence as a subject in the enunciating present assures that the past participle “TORN” can modify the pieces of his flesh without fully annihilating his subjectivity; otherwise, he could not be speaking. On the contrary, his past subjectivity as a member of the neighborhood is not secured in reality but must be insisted upon through the reframing of an absence, “TORN PIECES”, as a presence — a point of identity he longs to retain. Once again addressing the neighborhood in the second-person, he concludes, “Therefore, you [the neighborhood] ARE”, letting the copula emphatically pronounce an existence that is not fleshy—not embodied (59).

The dilemma the speaker alludes to between a subject whose capacity to enunciate hinges on its embodiment and a subject belonging to a recollected unembodied past is the problem of collective identity. The speaker seeks to identify with the collective of the neighborhood, but all he can guarantee for himself is his own individual identity. But this does not mean that the unbridgeable distance between a representation of the past and a present reality makes the effort of recollection fruitless for the speaker. On the contrary, to inhabit a collective identity while remaining oneself as a necessary feature of being an individual already implies a reconciliation with that distance. The speaker marks his acknowledgement of this fact by not insisting that his desire

to belong to a collective was ever wholly fulfilled: “i needed you then... identity... a sense of belonging / i need you now” (59). If, in the poetic present, the closest he comes to fulfilling this need to have a sense of belonging is through the recollections of La Loma he conjures up, the implication is that collective identity had only been secured for him in the past by insisting that, despite (or maybe because of) the necessary distance between the two, representations trump reality. It does not matter for him that the sense of La Loma as a place he has been conveying in the poem is at best only partial, for the selection of certain representations—those salient to barrio culture—was, in the final analysis, already arbitrary. But “arbitrary” does not mean meaningless.

Nor does the arbitrariness of representation preclude its potential for identity formation. Isolated from the barrio he once knew, the speaker insists that La Loma keeps him “away from INSANITY’S hungry jaws” (59). The speaker’s fear that insanity will consume his sense of self amounts to the claim that, for him, holding onto an individual identity involves maintaining a collective identity. Read through the analysis above on the speaker’s acknowledgement of the nature of representation, this claim does not entail that collective identity is independently real of the speaker and other individuals. Rather, it entails that maintaining a collective identity is an especially productive act of representation. Moreover, for the speaker, its productivity lies in part in the identifications it enables for him with people he does not know, and places he has not experienced. The poem’s final stanza consists of the speaker’s

“
Artworks achieve an
abstractionist effect by calling
attention to their own artificial
character in a move that invites
us to question not only the
aesthetic representations
proffered in art but the social
facts to which they allude.
”

laying claim first to La Loma and then laying claim to other Chicano neighborhoods through his repetition of the possessive pronoun “my” (60). He is not, however, staking this claim on being able to recollect anything about all the neighborhood he names, but on his identification with other people who maintain memories of their own neighborhood. The collective identity tying all of these distinct individuals together is being a Chicano from a barrio, and the process through which it comes about is the individuated production of aesthetic representations.

This notion of a collective identity is radically different from Ybarra-Frausto’s representationally realist take in “Rasquachismo.” Indeed, the abstractionist notion of identity is the inversion of the realist notion in the role it ascribes to representation. The latter subordinates representation to reality in its insistence that the barrio as basis of collective identity precedes individual acts of aesthetic representation. Only *after* drawing upon a pre-representational barrio culture can Chicano artists said to be participating in a rasquache aesthetic sensibility. In contrast, the abstractionist view of identity exemplified in “Un Trip” holds that individual acts of aesthetic representation together produce and maintain the reality of the barrio. The difference is important because it indicts Ybarra-Frausto’s descriptive aesthetics program as a subtle form of prescribing a politics to art. I do not intend to fully elaborate on this point here, but already from the difference between the realist and abstractionist notions of collective identity, one can see how, among other things, each view entails a position on authenticity. Briefly: if Chicano barrio culture exists in a pre-representational sense, then individual artworks may fail to be authentically Chicano to the degree that they do not accurately represent Chicano barrio culture. However, if that culture owes its existence to the myriad ways it has been represented, then, assuming that the artist sought to represent Chicano barrio culture, no representation is less authentically Chicano than any other. More generally, the abstractionist view entails a methodological, and political stance, where the focus on describing/interpreting Chicano art lies firmly on how individual artworks or aesthetic projects such as that of Raúl Salinas are participants in an always in-the-process formulation of a Chicano culture. Descriptive aesthetics affects nothing less than how one looks at art, and what one sees.

References

- [1] English, Darby. *How to See a Work of Art in Total Darkness*. The MIT Press, 2007.
- [2] Harper, Philip Brian. *Abstractionist Aesthetics: Artistic Form and Social Critique in African American Culture*. NYU Press, 2015.
- [3] Jakobson, Roman. “Two Aspects of Language and Two Types of Aphasic Disturbances.” In *Language and Literature*. Harvard University Press, 1987.
- [4] Noriega, Chon. “The Orphans of Modernism.” *Phantom Sightings: Art after the Chicano Movement*. University of California Press, 2008, pp. 16-43.
- [5] Raúl Salinas. “Un Trip through the Mind Jail.” *Un Trip through the Mind Jail y Otras Excursions*. Arte Público Press, 1999.
- [6] Saussure, Ferdinand de. *Course in General Linguistics*, trans. Wade Baskin. The Philosophical Library, 1959.
- [7] Ybarra-Frausto, Tomas. “The Chicano Movement/The Movement of Chicano Art.” *Exhibiting Cultures: The Poetics and Politics of Museum Display*, eds. Ivan Karp and Steven D. Levine. Washington and London: Smithsonian Institution Press, 1991, pp.

128-149.

[8] —. “Introduction to the First Edition.” *Un Trip through the Mind Jail y Otras Excursions*. [1980]. Arte Público Press, 1999.

[9] —. “Rasquachismo: A Chicano Sensibility.” *Chicano Art: Resistance and Affirmation*, eds. Richard Griswold Del Castillo, Teresa Mckenna, and Yvonne Yarbro-Bejarano, University of Arizona Press, 1991, pp. 155-62.

The Dilemma of Disempowerment: How Generational Trauma Impacts Alcohol Abuse Rates in the Australian Aboriginal Community

MAHIMA KRISHNAMOORTHY

STANFORD UNIVERSITY

The diverse aboriginal populations within Australia share an unfortunate link – significant health disparities relative to the non-indigenous population. Australian aboriginal communities not only have higher rates of physical illness, but also a high prevalence of mental illness, specifically addiction to drugs and alcohol. This paper will explore the causes behind this health inequality, as well as the current methods of treatment that are being implemented, and their cultural competency. This paper argues the necessity for evidence based biomedical strategies that work alongside indigenous care. The paper further asserts the importance of the comprehensive evaluations of these programs to ensure their longevity and prolonged success. The conclusions of this examination into substance abuse finds a prolonged history of devaluing, dominating and displacing a large group of people, which can push individuals, years later, to abuse alcohol and drugs. Specifically, the connection between personal and structural violence is made within the analysis, demonstrating that a potential connecting force exists between different Aboriginal groups, with important intergenerational ramifications.

Introduction

Anthropologist and physician, Paul Farmer, discussed global health inequity in one of his most significant works, *On Suffering and Structural Violence*, writing that “it is one thing to make sense of extreme suffering - a universal activity, surely - and quite another to explain it” (Farmer 1996). First coined in 1969 by Johan Galtung, structural violence refers to “violence exerted systematically by everyone who belongs to a certain ‘social order’ onto those belonging to a different class” (Farmer: 1996). Here, violence is used atypically, including not only physical abuse but also economic, social and emotional violence. According to the principle, attitudes and actions of certain people throughout history, including oppression, disempowerment and displacement contribute to systemic issues within the social infrastructure that continue to impact minority groups, long after the direct actions of disempowerment have been resolved.

This significant theory can be applied to various disenfranchised bodies around the world, especially in Aboriginal populations of Australia. The Aboriginal, or First Nation’s, communities experienced mass displacement and disempowerment when European settlement in Australia began in the late 18th century. While the relationship between indigenous and non-indigenous populations has certainly improved, structural inequalities still exist within these groups, such as education access, employment inequity and especially health disparity. Indigenous populations have higher rates than their non-indigenous counterparts of diabetes, cardiovascular disease, chronic respiratory disease, and mental illness including alcohol and drug addiction (Gracey et al. 2009). High rates of drug abuse and alcohol abuse have been noted in the indigenous communities, with some 95% of urban Aboriginal and Torres Strait Islander populations considering alcohol abuse as a serious problem in their community (Ministerial Council on Drug Strategy 2001). This

paper will focus on high substance abuse rates among Aboriginal populations, answering two main questions: (1) how are current interventions being implemented to alleviate this health inequality and (2) what are the main causes behind this phenomenon?

Aboriginal populations/communities will be used as an overarching term that describes most indigenous groups in Australia. This, of course, is a major over-simplification as each indigenous group lives on different land, speaks unique languages, and practices distinct cultural activities and ceremonies. However, given that the focus of the study is generational trauma as a potential connecting force between all of these groups, and that there is a dearth in published and peer-reviewed evidence on specific indigenous groups, the paper aims to understand the issues of health inequalities in indigenous populations, with some use of appropriate generalizations.

This paper will discuss structural violence and generational trauma in the aboriginal community, and its impacts on the relatively high rates of substance abuse. First, the paper will assess social determinants and disease patterns of indigenous health, considering the interactions between social and political variables. The paper will then aim to understand the Aboriginal notions of health and healing, and how culturally appropriate (or inappropriate) public health interventions have been used by the government to alleviate these structural health inequalities in alcohol and drug abuse. Specifically, this paper aims to make a distinction between personal and structural violence in indigenous populations, terms that will be further defined later, using substance abuse as a lens for health inequalities. Moreover, the paper will attempt to understand the relationship between these types of violence, and if one type of violence presupposes the manifestation, or latent, presence of the other.

Aboriginal Health Disparities

A clear divide exists in health status between indigenous and non-indigenous groups. Indigenous Australians consist of 3.3% of Australia's population; however and social problems is disproportionately higher than their non-indigenous counterparts. Australian Aboriginal people have a startlingly low life expectancy at birth, about 11 years less than that experienced by other Australians (AIHW). The leading causes of death for Aboriginal Australians demonstrate an interesting divergence between indigenous and non-indigenous populations, specifically within the "external causes" group, which the Australian Institute for Health and Wellness describes as including vehicle accidents, violence, drug abuse, suicides, and essentially all causes of death outside of the biomedical sphere. These external causes often are related to alcohol and drug abuse, illustrating their significance in indigenous communities. In fact, in a comprehensive study on mental illness in Aboriginal communities, 43% of respondents received a diagnosis of at least 1 DSM (Diagnostic Statistics Manual)-III disorder and 18.4% displayed substance use disorders, with alcohol dependence at 9.2%. The 2010 National Drug Strategy Household Survey showed that these rates were higher than the general Australian population, in which Aboriginal people were almost twice as likely to be recent users of illicit drugs as other Australians (Government 2011). Even more so, rates of comorbidity, or the presence of a secondary-linked disease, were very high, with almost one-half of those with behaviour or affective disorders meeting criteria for a substance use disorder (Kirmayer 2000). In 2011, alcohol accounted for an estimated 8.3% of the overall burden of disease among Aboriginal and Torres Strait Islander Australians – a rate 2.3 times higher than among non-Indigenous people (Government 2011).

In addition to harms and health, high levels of alcohol use can contribute to a range of social harms. Aboriginal people remain the most frequently arrested and incarcerated group in Australia; a 2016 survey noted that the percentage of Aboriginal prisoners was 27.6%, a number much higher than the total proportion of Aboriginal people in Australia. Sixty five percent of Indigenous homicides involve both the victim and offender having consumed alcohol at the time, which is three times more than the occurrence for others (Georgatos 2013). The aforementioned statistics demonstrate the significance of alcohol on the indigenous community in Australia, contributing to societal and biomedical disadvantages.

The historical context of alcohol and Aboriginal populations is important to understand when discussing how the issue presents itself today. Many social theorists propose that rates of alcohol use in Aboriginal populations began to rise due to European colonization, when the First Fleet arrived and pubs opened on the continent for the first time. Drinking was a prominent part of colonial life, and exposure to the act was likely the most significant influence on Aboriginal life, impacting emerging drinking patterns. In fact, many Aboriginal laborers were actually paid in alcohol or tobacco. However, in the 19th century, under the realization that Aboriginal and Torres Strait Islander population was dying out, colonial and, later, state and territory governments adopted policies "protecting" Aboriginal populations and thus implemented a forced prohibition for Aboriginal people. These assimilation policies rested on the paternalistic assumption that Aboriginal people could not control their own destinies and that the government needed to step in to "aid them". The Licensed Publicans Act of 1838 in New South Wales was the first legislation to forbid the supply of alcohol to Aboriginal populations in Australia, with all mainland colonies subsequently

following suit (McCorquodale 1985). This led to increased rates of illegal purchases of alcohol, more Aboriginal people put in jails, and a higher proportion of the community participating in "binge drinking" (Ministerial Council on Drug Strategy 2001). These barriers to alcohol had long been promoted as a method of colonial governments to claim to be "responsible" for the welfare of indigenous peoples. The final prohibition legislation was repealed in 1964, allowing Aboriginal populations access to alcohol, which was considered to be a major civil rights achievement linked to the equality and autonomy of Aboriginal peoples. Unfortunately, the perspective of Aboriginal Australians being unable to control their alcohol intake has persisted throughout time. A field survey of mostly white Australians at Mt Isa, Queensland reported a stereotyped view that Aboriginals "always get drunk" and "drink three times as much as us" (Khan 1990). In fact, the proportion of Aboriginal Australians that drink any alcohol is much lower than that same proportion of non-indigenous groups. However, it is the percentage that does choose to drink that does so at levels harmful to their health, with some studies placing that proportion at 68% of current Aboriginal drinkers.

Fortunately, state governments have been able to repeal most of the discriminatory alcohol legislations that were unfairly targeting Aboriginal peoples. In 2017, the Northern Territory government repealed two policies introduced four years prior: alcohol prevention orders (APOs) and alcohol mandatory treatment orders (AMTs). The former allowed police to "issue an order to a person charged with an offense if they believed that person was affected by alcohol at the time" essentially criminalizing alcoholism with 86% of people issued orders being Indigenous (Davidson 2017). The latter policy allowed police to issue orders to anyone arrested for intoxication three times in two months causing them to be placed in forced three months of treatment while incarcerated (Davidson 2017). It will be certainly interesting to see how legal developments will affect incarceration rates of Aboriginal people, and whether they will decrease the likelihood of an Aboriginal person to be arrested, especially as culturally appropriate interventions continue to be on the rise.

Notions of Health and Healing

Psychiatrist and anthropologist Arthur Kleinman theorized about a patient's "explanatory models", which he described as the perception and understanding of the causes, mechanisms, symptoms and effects of treatment surrounding their disease. In fact, within medical anthropology, disease and illness are considered as two different terms with completely distinct cultural significance. While disease is proposed to be biologically and scientifically determined, illness is considered to be culturally shaped, based on how each person perceives, experiences and copes with the disease. These perceptions influence the explanations of sickness, or Kleinman's explanatory model. For years, the understanding of indigenous health has largely been focused on a Western or European explanatory model. Alternatively, Indigenous peoples define wellbeing and health as being much more than just the absence of disease. For example, the Anishinabek (Ojibway) word *mno bmaadis*, "which translates into living the good life or being alive well, encapsulates beliefs in the importance of balance. For many other Indigenous peoples, land, food, and health are key components of being alive well (King 2009). Thus the Indigenous definition of health caters to many more concepts than the biomedical definition does. While some might

see this as unnecessary, it is interesting to note that even the WHO adapted their original definition of health, “the absence of disease”, to “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” (Constitution of the WHO). This recognition of multiple components as essential to one’s health is important to Indigenous notions of healing, and is necessary for developing culturally appropriate interventions.

Culturally Competent Interventions

It is clear that medical interventions are not “one size fits all”, and notions of health and healing are incredibly important in creating culturally appropriate and competent forms of treatment processes. Some interventions have been less successful than others. Most infamously are communities of indigenous populations that have been organized as alcohol-free, dry-camps in the bush. These approaches have oftentimes been undercut by community members secretly bringing alcohol into the camps, or even concocting their own alcoholic beverages. Most researchers agree that these programs largely did more harm than good, pushing people further into the fringes of communities and incarcerating many others (Davey 2015).

In his article *Culturally Appropriate Means and Ends of Counseling*, Rod McCormick states that “effective healing for First Nations people focuses on interconnectedness rather than autonomy, which is a more common goal for Western therapy” (McCormick 1995). McCormick goes on to note that that aim of healing for many indigenous people is to maintain balance between four dimensions: “physical, mental, emotional, and spiritual” (McCormick 1995). Some might argue that conventional biomedicine weights the physical or biological aspects of health over everything else. This might be a reason why conventional biomedical intervention strategies are typically not successful in the long-term for Australian minority populations, as the “universally accepted notion is that treatment and rehabilitation for native people should be culturally appropriate” (Brady 1995). Most programs at the turn of the 21st century were varying adaptations of the 2 step program of Alcoholics Anonymous, which represents a strictly biomedical and conventional approach to treatment without any room for discrepancies based on the target demographic.

Fortunately, the trajectory of this field has pushed addictions programs towards cultural competence and innovation. For example, the Ngarlu model of mental health counseling for Aboriginal Australians is highly regarded as a culturally appropriate methodology, using the Karajarri word, Ngarlu, for defining the place of the inner spirit (King 2009). This model allows participants to assess how their alcohol and other drug use affects their Inner Spirit and their connections to family, community, and country. The program has recently become particularly important for the Government of Western Australia’s Drug and Alcoholic Office. One example of a particularly simple culturally competent awareness campaign was a move towards encouraging Aboriginal women, given that most research indicates that they rarely drink at harmful levels, to play an important role in improving rates of alcohol abuse in their communities (Danieli 2010). Another example of good practice within addiction treatment is creating culturally secure facilities that are staffed by Aboriginal workers. This development has been shown to have positive effects on attracting and retaining Aboriginal people in treatment. Together, these types of programs and developments in the field of addiction treatment for Aboriginal people not only address the alcohol and drug problems of patients

but also the underlying the social and historical factors that affect the health and wellbeing of Aboriginal people. The move towards cultural competence has been extremely important to the field and provides hope for much more successful interventions in the future that account for alternative notions of health and healing, like those aforementioned.

One of the most comprehensive programs is the Queensland Indigenous Alcohol Diversion Program (QIADP), which began in June 2007 and is described as “indigenous specific”. The program is essentially used as a secondary punishment option for Aboriginal people that have been arrested for an alcohol related offense. The magistrate decides if the defendant can be considered for QIADP and an initial screening assesses the individual’s suitability to the program. While the assessment is conducted, some of the defendants are placed on bail, but in many cases, the defendant is incarcerated until a decision is made. Once the magistrate does endorse the treatment plan for the defendant, however, the defendant is bailed into the program and becomes a QIADP participant. The program itself runs for 20 weeks with regular follow-up calls and Progress Reports. Even more so, Aftercare programs that help participants reintegrate better into mainstream society is available to individuals who choose to participate after their treatment has completed. QIADP has been implemented in Cairns, Townsville and Rockhampton, and an extensive post-evaluation survey conducted by the Queensland Government found that overall there was a significant decrease in alcohol-related offenses before the program and after the program was completed. However, analyzing the success of these programs, all of the sites, except Cairns, observed an increase in the frequency of alcohol-related offending from the during-QIADP period to the post-QIADP period. This certainly doesn’t take away from the success of this program, but as the evaluation stated “this increase in alcohol-related offenses after the program completes is worthy of further research” (QIADP).

Unfortunately, while QIADP underwent an extensive post-evaluation survey, not many alcohol treatment programs that have been implemented enjoy the same luxury. In a 2010 comprehensive study of alcohol and smoking interventions for Aboriginal people, Clifford et al. discovered that although 18 of the 20 interventions studied utilized indigenous involvement in the planning and implementing of the program, only seven reported any indigenous involvement in the evaluation of the program. As the WHO explains “monitoring and evaluation of any alcohol program or intervention is vital to determine whether it works, to help refine programme delivery and to provide evidence for continuing support of the programme” (WHO 2007). While the trajectory of culturally appropriate alcohol intervention programs has brought much more indigenous involvement in implementation, the lack of Aboriginal voices in the evaluation of these programs is a stark issue that needs to be addressed. Often, it is not always the fault of the program creators that the evaluation process is not as successful as it should be. “Intervention evaluation research is complex, requiring resources, expertise and skills unlikely to be available in local Indigenous communities” (Clifford 2013). Because of this, many programs that exist today do not have any evidence-based protocols, which can often lead to ineffective interventions. Even more so, the fewer programs that have strong evidence supporting their methodologies, the less data that is compiled that can be applied to future interventions. Evaluation research needs to be prioritized in the field of substance abuse treatment programs, not only for the intervention community

itself, but also for state and federal governments. This is the next step in the continued trajectory towards culturally appropriate health programs.

Generational Trauma

Philosopher and social theorist Michel Foucault's wrote in his significant essay *Nietzsche, Genealogy, and History*, "Humanity does not gradually progress from combat to combat until it arrives at universal reciprocity, where the rule of law finally replaces warfare; humanity installs each of its violence in a system of rules and thus proceeds from domination to domination" (Foucault 1978). Farmer and Galtung would progress this analysis of domination one step further, and emphasize the importance of structural violence, and its compound effect on those being dominated, and the future generations. This concept of generational trauma, passed down throughout the years, has often been described in indigenous populations around the world as a major force that influences structures of inequality. "The high rates of suicide, alcoholism, and violence, and the pervasive demoralization seen in Aboriginal communities, can be readily understood as the direct consequences of a history of dislocations and the disruption of traditional subsistence patterns and connection to the land" (Kirmayer 2000). This theory does not only apply to Aboriginal communities in Australia. It instead makes the claim that the disempowerment and displacement of indigenous populations during European colonization has placed current indigenous groups into a tangled web of institutionalized injustice, unconscious prejudice and biopolitical disadvantage.

Contemporary institutions of structural violence can be observed as higher unemployment rates for Aboriginal people and worse access to higher educational opportunities (citation), leading many Aboriginal people to turn to substance abuse, often as a form of a coping with the systemic inequalities they experience. This generational trauma and structural violence connection has been documented as a strong force within tribal communities. For elders, the issues of native title to land and reinforcing cultural practices on country have been, of course, extremely difficult since colonization. Stresses caused by loss of land and loss of culture have led to increased rates of depression in modern tribal Elders, whose role in their communities has been eroded by these factors (Jimenez 2012). These stresses held by elders can cause defragmentation within a tribal community, which of course, can have impacts on the youths within the community. As Kirmayer explains, "Cultural discontinuity has been linked to high rates of depression, alcoholism, suicide, and violence in many communities, with the most profound impact on youth" (Kirmayer 2000). Young Indigenous Australians are more than twice as likely as their non-Indigenous counterparts to die from alcohol-attributable causes (Chikritzhs 2004). This relationship between elders and younger members of the community is incredibly important, but structural violence has been inflicted upon these relationships in the past and present. Hunter (1994) commented on the significance of male figures in young Aboriginal lives, and how this connection has been fragmented by familial separation, along with high incarceration rates. Farmer could see this

“
Evaluation research must
be considered a priority in
intervention programs in order to
provide effective evidence-based
programs for Aboriginal persons
”

as a contemporary form of structural violence, a systemic issue that percolates throughout the community and impacts many generations of Aboriginal people. Hunter, without using the term of structural violence, seems to agree with this sentiment, pointing out that "early mortality and excess morbidity from alcohol-related causes, enormous rates of arrest and detention, absence from communities and families in pursuit of alcohol, and the dysfunctionality of intoxication, all disproportionately impact the availability of males as parents" (Hunter 1994).

Traumatized people often use alcohol and other drugs as forms of self-medication and as a means of coping with feelings (Robin et al. 1996). Often times, alcohol and drug abuse lead to increased rates of violence, another potential expression of pent up frustration and despair. An ethnographic study of the transgenerational effects of alcohol, demonstrated that many of the men in the study group expressed rage both in private (e.g. against family members) and public (e.g. street violence) contexts. In all instances of violence, the men were under the influence of alcohol or drugs (Atkinson 2002). This analysis of alcohol abuse rates as a "symptom" of the problem of generational trauma can, unfortunately, be seen in indigenous groups beyond those in Australia. In Maori populations in New Zealand, indigenous groups in Canada, and Native Americans in the United States, high-rates of alcohol-related suicide and alcohol abuse is observed, some with the highest rates out of any ethnic group in their country. These statistics demonstrate a potential connecting force between these groups, and this paper reaches the conclusion that the generational trauma and contemporary structural violence are those connecting forces.

Conclusion

Using substance abuse as a proxy for understanding health inequalities between indigenous and non-indigenous groups in Australia delineates more than just statistics on health. It outlines the importance of culture and societal issues on individual biomedical health. This paper makes the claim that the dislocation and disempowerment of Aboriginal Australians can be understood as structural violence, impacting the systemic issues that have filtered down to have indirect effects on individual Aboriginal peoples. Structural violence could even be seen as the larger "disease", with alcohol abuse as a symptom. This representation of the health issue demonstrates a potential connecting force between Aboriginal peoples, providing reasoning behind this vast gap in health inequality and substance abuse rates between them and their non-indigenous counterparts.

This paper also concludes that the trajectory of intervention programs, while moving towards cultural competence, still has much more to do. Evaluation research must be considered a priority in intervention programs in order to provide effective evidence-based programs for Aboriginal persons. Evaluation strategies that emphasize continuous quality improvement throughout different stages of the intervention have been found to be very important for achieving positive outcomes for the programs (Wandersman et al., 1998, Dusenbury & Falco, 1995; Hansen, 2002). As evaluations are time-consuming and costly, it is absolutely necessary for the state and

federal government to make statements and even policy changes that enforce the requisite of evaluations which can increase opportunity for funding and grants in this field. Moreover, evaluations must include indigenous involvement in order to ensure that the analysis of the success of a program includes the voices of those that actually participated in it. Programs should have a clear outline of goals for their intervention that can be compared to documented results and post-intervention surveys. What seems clear is that any intervention that is implemented should be a blend of evidence-based biomedical and indigenous care strategies. Moreover, enactments of these interventions should be done with cognizance of power dynamics, understanding of alternate notions of healing and health, awareness of the historical contexts of suffering, and a familiarity with the interconnectedness of beliefs.

The reality is that the pursuit of autonomy, and biological, biosocial, and biopolitical healing is “in the end, an ongoing intergenerational struggle to define and redefine and practice what is wellness” (Million 107). There is no set rulebook for how to heal from and reach a healthy status following centuries of dislocation. Recognizing that Aboriginal culture is not just something of the past, but rather an ongoing and adapting collection of ideas and ways of living, is of utmost importance to ensure that programs are culturally cognizant.

References

- [1] Alcohol in Australia, Issues and Strategies. (2001). Ministerial Council on Drug Strategy.
- [2] Atkinson, J. (2002). *Trauma trails: Recreating song lines: The transgenerational effects of trauma in Indigenous Australia*. North Melbourne, Vic.: Spinifex Press.
- [3] Australian Institute for Health and Wellness. (2018). *Deaths in Australia*. Australian Institute for Health and Wellness. Web Report.
- [4] Brady, M. (1995). *Culture in treatment, culture as treatment. A critical appraisal of developments in addictions programs for indigenous North Americans and Australians*. Great Britain. Elsevier Science.
- [5] Chikritzhs, T., Pascal, R. (2004). *Trends in Youth Alcohol Consumption and Related Harms in Australian Jurisdictions, 1990-2002*. Bulletin No. 6. National Drug Research Institute, Curtin University
- [6] Clifford, A., Pulver, L. J., Richmond, R., Shakeshaft, A., & Ivers, R. (2010). Smoking, nutrition, alcohol and physical activity interventions targeting Indigenous Australians: Rigorous evaluations and new directions needed. *Australian and New Zealand Journal of Public Health*, 35(1), 38-46. doi:10.1111/j.1753-6405.2010.00631.x
- WHO (2007).
- [7] Constitution of the world health organization. Basic Documents, Forty-fifth edition.
- [8] Danieli, Y. (2010). *International handbook of multigenerational legacies of trauma*. New York: Plenum Press.
- [9] Davey, M. (2015, February 20). Alcohol abuse behind high rates of early death among Indigenous, study finds. Retrieved from <https://www.theguardian.com/australia-news/2015/feb/20/alcohol-abuse-behind-high-rates-of-early-death-among-indigenous-study-find>
- [10] Davidson, H. (2017, August 18). Northern Territory repeals alcohol laws ‘discriminatory’ to Indigenous people. Retrieved from <https://www.theguardian.com/australia-news/2017/aug/18/northern-territory-repeals-alcohol-laws-discriminatory-to-indigenous-people>
- [11] Dusenbury, L., & Falco, M. (1995). Eleven components of effective drug abuse prevention curricula. *Journal of School Health*, 65, 420 – 425
- [12] Farmer, P. (2009). On Suffering and Structural Violence: A View from Below. *Race/Ethnicity: Multidisciplinary Global Contexts*, 3(1), 11-28.
- [13] Galtung, J. (1969). Violence, Peace, and Peace Research. *Journal of Peace Research*, 6(3), 167-191. doi:10.1177/002234336900600301
- [14] Gracey, M., & King, M. (2009). Indigenous health part 1: Determinants and disease patterns. *The Lancet*, 374(9683), 65-75. doi:10.1016/s0140-6736(09)60914-4
- [15] Georgatos, G. (2013). Australian Bureau of Statistics on prison rates. Retrieved from <http://thestringer.com.au/australian-bureau-of-statistics-on-prison-rates-1928#.XAUrehMzY1h>
- [16] Government of Western Australia Drug and Alcohol Office (2011). *Aboriginal Drug and Alcohol Framework for Western Australia*
- [17] Hansen, W. B. (2002). Program evaluation strategies for substance abuse prevention. *Journal of Primary Prevention*, 22, 409 – 436
- [18] Hunter, E. (1999). *Aboriginal health and history: Power and prejudice in remote Australia*. Cambridge: Cambridge University Press.
- [19] Jimenez, D., Bartels, S., Cardenas V., Daliwal, S., Alegria, M. (2012). Cultural beliefs and mental health treatment preferences of ethnically diverse older adult consumers in primary care. *American Journal of Geriatric Psychiatry*
- [20] Kahn, M., Hunter, E., Heather, N., & Tebbutt, J. (1991). Australian Aborigines and alcohol: A review. *Drug and Alcohol Review*, 10(4). doi:10.1080/09595239100185411
- [21] King, M., Smith, A., & Gracey, M. (2009). Indigenous health part 2: The underlying causes of the health gap. *The Lancet*, 374(9683). doi:10.1016/s0140-6736(09)60827-8
- [22] Kirmayer, L. J., Brass, G. M., & Tait, C. L. (2000). The Mental Health of Aboriginal Peoples: Transformations of Identity and Community. *The Canadian Journal of Psychiatry*, 45(7), 607-616. doi:10.1177/070674370004500702
- [23] Kunitz SJ, Gabriel KR, Levy JE, Henderson E, Lampert K, McCloskey J, and others. Alcohol dependence and conduct disorder among Navajo Indians. *Journal of Studies on Alcohol* 1999;60(2): 159-67
- [24] Kunitz, SJ, *Disease and Social Diversity: The European Impact on the health of Non-Europeans*, Oxford University Press, New York (1994)
- [25] McCormick, R. *Int J Adv Counselling* (1995) 18: 163. <https://doi.org/10.1007/BF01407960>
- [26] McCorquodale, J C., (1985). *Aborigines, a history of law and injustice*. University of New England.
- [27] Million, Dian. (2013). *Therapeutic Nations: Healing in an Age of Indigenous Human Rights*. University of Arizona Press
- [28] Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., sMorrissey-Kane, E., & Davino, K. (2003). What works in prevention: Principles of effective prevention programs. *American Psychologist*, 58(6-7), 449-456. doi:10.1037/0003-066x.58.6.7.449m
- [29] Saggars, S., & Gray, D. (2009). *Dealing with alcohol indigenous usage in Australia, New Zealand and Canada*. Cambridge: Cambridge University Press.

- [30] Queensland Police Service. (2010). Queensland Indigenous Alcohol Diversion Program (QIADP) Recidivism Study. Specialist Courts and Diversion, Legal Services Branch.
- [31] Robin, R.W., Chester, B., and Goldman, D. (1996) "Cumulative Trauma and PTSD in American Indian Communities" in *Ethnocultural Aspects of Post Traumatic Stress Disorder: Issues, Research and Clinical Applications*.
- [32] Wandersman, A., Morrissey, E., Davino, K., Seybolt, D., Crusto, C., Nation, M., et al. (1998). Comprehensive quality programming: The eight essential steps to effective community-directed prevention programs. *Journal of Primary Prevention*, 19, 3–31.
- [33] Weatherburn, D. (2014). *Arresting Incarceration: Pathways out of Indigenous imprisonment*. Aboriginal Studies Press.
- [34] *Working together: Aboriginal and Torres Strait Islander mental health and wellbeing principles and practice*. (2014). Canberra: Department of The Prime Minister and Cabinet.

The Spirit of Antitrust: How Antitrust Lost Its Way, and How It Can Be Revived

NOAH NAPARST

JOHNS HOPKINS UNIVERSITY

This paper explores the history of antitrust in the United States, and how the goals of competition policy evolved over time. In particular the paper explains how antitrust was enfeebled in the post-war, and how this has contributed to current economic issues. The link between monopolies and fascism is also discussed.

Introduction

“The problems with which the antitrust laws are concerned—the problems of distribution of power within society—are second only the questions of survival in the face of threats of nuclear weapons.” Thus, warned Lee Loevinger, President John F. Kennedy’s antitrust czar, during his interview for the role. He spoke of antitrust as his “secular religion,” an antidote to the excesses and evils that accompany industrial bigness. Senator John Sherman, the namesake of the first antitrust legislation in the United States, warned of the unchecked powers possessed by amassers of enormous amounts of capital. To him, the “kingly prerogative” of the Gilded Age’s “robber barons” ran counter to the democratic ideals of the republic, and it was with this vision that the antitrust laws were first enacted and enforced in the United States.

Antitrust laws play a vital role in securing a competitive economy. They are used to break up or prevent excessively large firms that can use their market dominance to charge excessive prices, prevent innovation, and constrain individual and businesses’ economic freedom. They also prevent the accumulations of industrial power in the hands of a few decision makers, which serves a further political role, for antitrust laws serve as a complement to the democratic system of checks and balances. They serve as a constitutional restraint on private power: “the same principles that prevent Congress from delegating regulatory power to private entities—a restriction contained in constitutional law—also prevent private entities from taking regulatory power for themselves—a restriction contained in the antitrust laws.” Their role in protecting competition has been compared to that of the Bill of Rights; after all, “Competition is a public good, and society cannot expect the victims of anticompetitive conduct to protect themselves.” This is called “political antitrust”, and is one of the most important functions of antitrust law that will be explored in this paper.

For the first seven decades of their existence, these laws ensured a “decentralization... of influence and power” that ensured a fairer, more democratic society. Countries like Germany that failed to reckon with their own nascent monopolies suffered from monopolies’ propensity to facilitate totalitarianism. Beginning in the 1970s, however, antitrust laws in the U.S. were devalued and stripped of their bite, to the detriment of the economy and American democracy. What were once viewed by their creators and enforcers as a constitutional check on private power became relegated to the

status of a mere “consumer welfare prescription.” What was the result? I need not regale you on the structural economic problems facing the U.S., but the troubling picture of a second Gilded Age emerges—in which 38.6 percent of the national wealth is owned by 1 percent of the population, where inequality, an unsteady economy, increased private sector concentration, and a political system ever more responsive to private money and less accountable to voter interests all threaten the health of our democracy. How did we get here? The current economic situation of the U.S. has been impacted by a variety of factors, to be sure, but the lack of antitrust enforcement has played a large part. The problems plaguing the United States in the 1890s and the problems plaguing it now face a common cause: increasing corporate consolidations. In order to begin tackling some of these larger issues, a dialogue about how these laws should be interpreted, and how they ought to be updated for the 21st Century, is in order. This paper will provide an outline of the history of antitrust in the United States. It will discuss how the interpretation of the antitrust laws, as well as understandings of their purpose, has changed. Lastly, it will argue for a return to robust antitrust enforcement, which ought to be a priority for those seeking reform of the U.S. political and economic system.

The Foundations of Antitrust

First, we must begin in 1890, with the passage of the Sherman Antitrust Act. It is remarkable in its brevity, containing just two sections that can be explained on the back on a napkin, and it is from this simplicity that the modern antitrust debate, and the current antitrust problem, arises. The law proscribes “every contract, combination in the form of trust or otherwise, or conspiracy in restraint of trade or commerce among the several states, or with foreign nations”, and likewise any act of monopolization or attempt at monopolization of said interstate commerce. That is all the act says; if taken literally it bans almost any possible commercial agreement. The interpretation of the law would therefore be up for debate in subsequent years, and the courts would have to make sense of what Congress had actually meant by its language when dealing with business practices like horizontal dealings, vertical arrangements, and the legality of dissolving trusts. The problem of deducing Congress’ intent in writing the Sherman Act is a difficult one to solve, but we can darkly observe that “Congress elected generally to leave specific enforcement decisions to the judiciary.”

The law was written with such broad language, forbidding so much, to appease those clamoring for government action against the increasing concentration of American industry without bringing about any radical action. Still, the need for antitrust law was there. A nation born in rebellion against mistreatment by British Crown monopolies, “a nation of founders and small-town entrepreneurs – ambitious, mobile, optimistic, speculative, anti-authoritarian, egalitarian, and competitive,” was beginning to resemble the America envisioned by advocates of *laissez-faire* capitalism in the Trust Movement – “centralized, run by great men, free from any government interference.” Between 1895 and 1904, the U.S. witnessed 2,274 manufacturing firms consolidate into 157 giants, the majority of which dominated their industries. The results were staggering: industrial titans, conglomerations of capital unprecedented in human history, largely controlled the economy. An owning class of 4,000 families controlled just as much wealth as everyone else combined while millions of farmers, industrial workers, and urban poor struggled. It really was a tale of two Americas: one an experiment in democracy and liberalism which had resulted in breakneck growth and the largest economy in the history of the world, and another an extremely unequal, crisis-prone economy controlled by a few private interests. If this picture seems familiar, it is because the same problems affecting the U.S. today share at least one similar root cause with the problems of the Gilded Age – the decline of competition. Consequently, just as the squalor of the cities, the depravities of child labor, and excesses of industrial capitalism provoked the ire of Progressive reformers, so did the trusts. Following a decade of unusually violent and frequent strikes, and with “wide discussion of alternative labor systems,” Congress decided to enact the Sherman Act to “head off direct regulation or Marxist solutions.” In the first decade of its existence the law saw little use. During the administration of Theodore Roosevelt, however, antitrust grew its teeth, and we must look to him when considering how to grow them back.

The Trust-Buster

Theodore Roosevelt is famous for many things, but one moniker of his that has endured is that of “trust-buster.” It is easy to imagine “old rough and ready” wrangling the out-of-control monopolies with the might of someone who was truly larger than life, but the reality of his crusade against the trusts is both more complex and more significant. It began just two weeks before the assassination of his predecessor, President McKinley, when he gave a speech condemning the trusts. He argued the “vast combinations of capital which have marked the development of the industrial system create new conditions, and necessitate a change from the old attitude of the State and the nation toward property.” Contrast this to the administration of President McKinley, which embodied *laissez-faire* in all but name. McKinley alluded to the trust problem just once in his presidency, in a state of the Union speech. Tellingly, when J.P. Morgan’s proposed buyout of Andrew Carnegie’s steel trust was announced, he held a White House dinner in his honor. The McKinley administration’s competition policy was “as if the Sherman Act did not exist.”

Just two weeks later, there was a President who was not just critical of the trusts, but supported government intervention against them, regardless of contemporary norms about the relationship between government and property. Roosevelt knew he had to proceed carefully and so he did. First, recognizing that without enforcement the Sherman Act was stillborn, he ordered

Attorney General Philander Knox to investigate Morgan’s Northern Securities Company, which was formed after a merger to monopoly in the railroad industry. In an unprecedented move, he announced that this trust violated the Sherman Act of 1890, and a few weeks later, his administration filed suit against the Northern Securities Company. Morgan’s attitude in this episode was telling, as he assured Roosevelt, “If we have done anything wrong, send your man to my man and they can fix it up.” Roosevelt later remarked that he felt as though Morgan saw his administration as a rival business, either to be bargained with or crushed. After two trials, the Northern Securities case made it to the Supreme Court, where Justice John Marshall Harlan sided with Roosevelt and provided the 5-4 majority needed to stop the merger and break up the trust. Roosevelt’s legal sally into antitrust enforcement was the beginning of an era of robust antitrust enforcement that would last well past his term. In planning his assaults on Northern Securities, and later other cartels, Roosevelt secured both legal victories allowing for later antitrust regulation and established a precedent of what was to be done about excessive bigness.

Roosevelt was a peculiar choice for a trust-buster. He was, after all, from the wealthy New York aristocracy himself, and admired the largeness and scale present in the new industrial behemoths, just as he respected size in affairs of state and institutions more generally. But he had a strong democratic impulse, and a sense of fairness that put the interests of the public before the interests of large firms. Further, while he was not anti-business and certainly no socialist, he had a clear understanding of public opinion. Roosevelt worried that without democratic reform of the excesses of industrial capitalism, the U.S. would be susceptible to the anarchist and communist unrest that was affecting Europe. For him, the trust problem was not really an inevitability of modern times, but a very real threat to the foundations of American democracy. It follows that the solutions to the trust problem would have to be centered on democratic ideals, and that the solutions would have to be cognizant of the constitutional relationship between private and public power. Roosevelt would have balked at the current view of antitrust laws as a “consumer welfare prescription,” to be used for punishing wicked behavior but not to effect a larger vision for a country’s political system. By Roosevelt’s logic, antitrust enforcement should not be seen as a technocratic affair, but instead as a political necessity. Supreme Court Justice Louis Brandeis, who excoriated “bigness” and also believed in antitrust’s political role, wrote, “Men are not free if dependent industrially on the arbitrary will of another.” As he put quite elegantly, “whether it has exceeded the point of greatest economic efficiency or not, [a corporation] may be too large to be tolerated among the people who desire to be free.” It follows that successful but abusive firms, or banks that are profitable but too big to fail, are problematic. It is from this basis that Roosevelt set out to wrangle the trusts, and it is this basis that has been forgotten by current politicians and Supreme Court justices. It is to these foundations that antitrust law and its interpretation must return, lest the U.S. slide deeper into its second Gilded Age.

The 1912 Antitrust Referendum

By 1912, antitrust law had become a “primary level of economic policymaking.” Roosevelt’s administration filed forty-five cases, with Taft adding another seventy-five in just one term, causing the breakup of almost every major trust in America. 1912 was also an election year – and it would be a pivotal one for antitrust policy. It

was the last election where debates surrounding antitrust played a decisive role, and “one of the few... where the public was engaged with and voting on what kind of economic order they wished to live in.” There were four candidates: Woodrow Wilson, William Taft, Theodore Roosevelt, and Eugene Debs. Each of them has endured in the American consciousness, with three serving as Presidents, and Debs mounting the most successful socialist Presidential campaign of the 20th Century. A modern analog might be Ronald Reagan, Barack Obama, Donald Trump, and Bernie Sanders running in the same Presidential election. Surely enough the contest was among the most dramatic in U.S. history, with the candidates trading insults, an assassination attempt on Roosevelt, the loss of Taft’s confidant Archie Butt in the sinking of the Titanic, and the death of Taft’s running mate days before the election.

At issue were not just the personalities involved, but a vision for the future of the economy. Each of the candidates proposed a different vision of the economy and thus the direction antitrust laws ought to take. The reasoning and politics behind their viewpoints are relatively complex, but Crane argues that they stem from three questions about how the American economy ought to work:

- (1) Do we want a competitive economy or a managed one?
- (2) Is antitrust necessary to a competitive economy?
- (3) What sort of institutional arrangements produce the best antitrust enforcement?

Both Roosevelt and Debs would have answered no to competitiveness on the first question. Roosevelt’s viewpoint might seem strange to those who know him as the famous “trust-buster”, but he had undergone somewhat of a change of heart over the course of his presidency. He had come to view his antitrust legacy as bringing “bad” trusts to heel, leaving the “good” trusts to be regulated not dissolved. In 1906, before initiating the suit against Standard Oil, Roosevelt proposed to its leadership that it accept government oversight, or even become a “public” trust. Standard Oil refused, and what could have been the American equivalent of Saudi Aramco instead was dismantled. Debs was also against industrial competition, albeit for different reasons. Recognizing the immense power of the industrial giants and adhering to a Marxist vision of human material progress, Debs thought the monopolies should not be dissolved, or regulated, but nationalized, their earnings collectivized. Further, he believed that more concentrated an industry, the easier it was to nationalize, and so did not believe in any antitrust enforcement whatsoever. Lastly, Debs faced added pressure to disavow antitrust from the socialist left, which criticized the law that was used to break up strikes and prevent unions from forming.

This left only Taft and Wilson as advocates for the kind of competition policy that had been practiced until that moment. While they both believed in using antitrust law as a way of ensuring competitiveness, they disagreed on the finer points of doing so. Taft favored what is called a common-law incrementalism approach, which leaves antitrust enforcement broadly in the hands of the court system. Wilson favored something in between what Taft and Roosevelt proposed, which stresses the importance of the courts in enforcing antitrust but empowers government agencies to investigate monopolists, as part of an expert-commission model. Wilson’s approach of an administration leading antitrust efforts promised more immediate action than Taft’s, whose approach depended on the antitrust leanings of the judges that were appointed, and the

Presidents who appointed them. Wilson’s credentials were bolstered because Taft, a Republican, was tied to his party’s support of import tariffs. As the other central economic locus of debate in that election cycle, import tariffs were seen as a tool used by the trusts to shield themselves from competition, and Wilson seized on this issue in a number of speeches. Wilson won the election, and it was his antitrust platform that was enacted. It came in the form of the Clayton Act of 1914 and the Federal Trade Commission, in addition to the reduction of tariffs.

The 1912 election was a clear referendum on the economic direction voters wanted the U.S. to take and serves as a reminder that almost no argument in antitrust is new. Current advocates of any competition policy ought to study this election for the lessons it provides. The relationship between the tariff and trust issues is worth noting, too. Voters should be skeptical of politicians who put up barriers to entry, and vigilant about who stands to gain from industry lobbying on trade. Lastly, the outcomes of the options championed by the other candidates, especially Debs’ and Roosevelt’s, are worth considering. Roosevelt’s regulated, but broadly anticompetitive industrial conglomerates would in Germany and Japan give rise to monopolies that contributed to the rise of fascism, and Debs’ calls for collectivized monopolies would be replicated in the Soviet Union. These experiments in what would later be known as corporatism, would have dire consequences.

I. G. Farben and the Nazi State: How Monopolies Give Rise to Fascism

The link between industrial monopolies and fascism has been explored before but there is an important lesson in the “bad history, bad policy, and bad law” that results from “[excluding] certain political values in interpreting the antitrust laws.” A useful example is the firm I. G. Farben in Germany. It began as Bayer in 1899, selling primarily aspirin, but would grow into the industrial arm of the Nazi regime. In 1904 Bayer formed a cartel with Afga and BASF, just as seven rival firms were doing the same. These two cartels merged in 1916, forming a “mega-cartel”, or profit-sharing pool, that coordinated on research, pricing, insurance, legal matters, and patents. Finally, in 1925, the members of the cartel consolidated into one “integrated, widely-held corporation,” completing the monopolization process. By the time that Hitler’s rearmament program was escalating, I.G. Farben became “deeply allied with and enmeshed with the Germany war effort.” The benefits to it were enormous: the firm extracted immense profits and grew its market share even as the rest of Germany continued to suffer from the Great Depression. By the invasion of Poland in 1939 it controlled nearly all chemical production, “including 100% of synthetic rubber, 100% of lubricating oils, 100% of serums, 90% of plastics, 88% of magnesium, 64% of explosives, and 75% of nitrogen.” The truly heinous result of I.G. Farben’s integration into the German war machine, however, was its use of slave labor and manufacturing of the Zyklon B gas used to exterminate millions in the Holocaust. While it was not the only German firm to participate in such atrocities, its highly centralized structure, which concentrated power in its president, made its role easier.

In a recent paper, Crane provides some historical reasons for I.G. Farben’s infamous rise, offering insights about the relationship between concentrated industries and the erosion of democracy more generally. It is important to note, after all, that what happened in Germany in the lead-up to World War II also occurred in pre-war

Japan and Italy. So, what is the link between these two phenomena?

First, the interests of monopolists and fascists overlap to a significant degree. Monopolies seeking higher profits can ally themselves with a regime which provides them with business; fascist regimes are particularly alluring because their programs of conquest provide a guarantee of sales. In addition to necessitating exploitable surges in government spending, wars of aggression provide new markets to dominate. This can be especially lucrative; in the territories annexed by Germany, I.G. Farben was the economic arm of conquest, “coercing firms in conquered territories to sell cheaply or simply taking them over.” Fascist governments benefit from consolidation, when firms become easier to coordinate with on military matters (as evidenced by the U.S.’s relaxing of antitrust efforts during the war), and easier to control. The Nazi government was heavily reliant upon I.G. Farben and other such conglomerates to enact its war mobilization. Hitler remarked in 1936 that “the minister of economics has only to set the tasks of the national economy; private industry has to fulfill them.” And while the state takes on control over key business decisions for its own purposes, the capital remains private. All the benefits of government economic policy are therefore privatized. During periods of popular unrest big business interests accept fascism over other insurgent movements like communism, of course because fascism allows them to remain private but also because they can profit from exploitation of the workforce. Thus programs of military buildup, wars of expansion, and the logistics of building a war machine suit the interests of both parties.

Second, monopolists and fascists can also aid and abet each other’s rise to dominance. The suspension of broad rule of law and constitutional checks allow authoritarian governments to favor certain firms and suppress labor movements with little political resistance. On the other hand, monopoly firms can use their abnormally bloated cash and sprawling investments to support fascist parties. On February 27, 1933 – the day of the infamous Reichstag fire – I.G. Farben deposited RM 400,000 into the Nazi Party’s reserves. Later the firm’s money was used to bankroll government projects, and its media resources were used to disseminate Nazi propaganda to bring other institutions under the control of the state. The institution of fascism in Germany has been described as “the political victory of... finance capital... coalesced by the end of 1932 on a policy bent on violent expansion and war.” Furthermore, large firms are more successful at wielding political power than small firms working together because of the nature of cartel organization, which is hampered by divergence on the finer points of political and economic interests.

Finally, the broader economic structure of a society is a unifying goal for monopolists and fascists. As mentioned earlier, the robber barons in the U.S. sought to create firms that were “centralized, run by great men, free from any government interference,” Simply replace “government interference” with “civil interference,” and the grand visions of both actors have much in common. Totalitarian and monopolistic systems lack the decentralized basis underpinning democracies and markets that allows “individuals in small groups... [to] strive for whatever they wish.” Instead both benefit from systems

that centralize power: fascism by strengthening the role of the state rather than constraining it, and in I.G. Farben’s case the series of steps to centralize power within the firm which allowed it to “[replicate] many of the democracy-quashing changes occurring in the political regime, with the effect of extending totalitarian control from the political to the business realm.” Their mutual interests, ability to aid each other’s rise, and, perhaps more abstractly, their common visions for society provide a clear platform for cooperation.

Deducing the direction of causality is trickier. Whether monopolies have historically caused movements toward fascism or whether fascism has resulted in dangerous levels of concentration has been a subject of academic debate. In Germany, however, the direction of causality is clear. The broad structure of the I.G. Farben monopoly preceded the rise of Nazism, with the major acts of consolidation in the chemical industry occurring decades before the ascension of the Nazi Party to power. Moreover, Hitler’s early

reliance on the firm’s “abnormal financial resources, ubiquitous local presence, and—in particular—the power to direct an entire industry” are all indicative of the preexistence of a monopoly. The case of I.G. Farben shows the U.S. made the correct decision in the 1912 election, eschewing an anticompetitive state of affairs that could have been dangerous had it been allowed to continue. More broadly, this episode shows that concentrated economic power can cause concentration of political power, which Crane argues has “corrosive effects on democracy.”

Governments and the public should therefore be watchful of trends in concentration, not just for their consumer welfare implications but also for their effects on politics. In the era of fascism-lite, there may not be political effects that rise to the magnitude of the horrors witnessed in Nazi Germany. Instead, one can look to subtler consequences of concentration, like money in politics (re: Citizens United), the use of consolidated media to bolster political campaigns (re: Roger Ailes’ support of Donald Trump or Silvio Berlusconi’s use of his media empire to dominate Italian elections), or the role that industrial firms have played in starting and prolonging U.S. military campaigns (re: Haliburton with the invasion of Iraq). All represent the erosion of democratic principles even if they do not amount to outright fascism and could be warnings of more sinister developments.

Heeding the lessons learned from World War II, American and European lawmakers drafted new antitrust legislation for the post-war age that consciously embodied anti-fascist impulses. In 1950, the U.S. Congress created the Celler-Kefauver Act of 1950, which focused especially on anticompetitive mergers. It closed loopholes that had allowed some firms to merge even despite the Clayton Act of 1914, but most importantly the law factored in clearly articulated pro-democracy ideals. In particular section 7 of the Act was written with Congress’ “clear concern that an economic order dominated by a few corporate giants could, during a time of domestic stress or disorder, facilitate the overthrow of democratic institutions and the installation of a totalitarian regime.” Senators Emanuel Celler and Estes Kefauver, the law’s cosponsors, warned of the extreme danger posed to democracy by industrial concentration, and in floor speeches, both cited Nazi Germany as a warning. The

“
Heeding the lessons learned
from World War II, American and
European lawmakers drafted
new antitrust legislation for the
post-war age that consciously
embodied anti-fascist impulses
”

Supreme Court even acknowledged Congress' intent in passing the law not just to challenge "accelerated concentration of economic power on economic grounds, but also of the threat to other values a trend toward concentration was thought to pose." The U.S. thus entered the postwar period with a clear, political stance on antitrust, updated with anti-totalitarian language. This moral stance was strengthened with tools to prevent mergers to monopoly before they happened rather than breaking up monopolies years after they formed. It would not be until the 1960s that the Chicago school of antitrust would succeed in convincing large swathes of economists – and the Supreme Court – that monopolies can be tolerated.

The Chicago School and the "Economics" of Antitrust

"Congress intended the courts to implement ... only that value we would today call consumer welfare. To put it another way, the policy the courts were intended to apply is the maximization of wealth or consumer ... satisfaction." This is the thesis of Robert Bork's 1966 paper, "Legislative Intent and the Policy of the Sherman Act," which Wu calls "the most influential single antitrust paper in history." This paper became a part of the cannon of the Chicago school of antitrust; its name refers to the intellectual developments in economics and other fields that took place in and around the University of Chicago from the 1950s until the 1980s. Bork himself innovated on the work of his mentor, Aaron Director, who worked with classic price theories to attack the antitrust case law as lacking concern about consumer welfare. But he took his mentor's message a step further. Bork argued not just about what contemporary antitrust law ought to do, but alleged that "consumer welfare", essentially competitive prices, is what the laws were solely intended to protect.

With a proverbial stroke of the pen, Bork succeeded in erasing six decades of democratic choice about what antitrust meant, in what has been called the greatest victory of the Chicago intellectual school. This reorientation of the intent behind the Sherman, Clayton, and Celler-Kefauver Acts, as well as abandonment of the legal conclusions in previous antitrust case law, has almost no basis in the historical record. But still by the end of the 1970s the Supreme Court had adopted many of Bork's principles, and by 1979 Chief Justice Burger was writing that "Congress designed the Sherman Act as a 'consumer welfare prescription,'" citing Bork's book. Why did his ideas gain such prominence?

To understand the Chicago school's rise from the fringes of the antitrust debate to prominence, it is important to understand the background of this era. The 1950s and 60s were the time of the Warren Court, which was arguably the greatest expansion of judicial power in American history. While it succeeded in expanding in civil liberties and improving racial equality, the Warren Court revolution became a victim of its own success. During this time, "judicial activism," or the courts acting in place of Congress to enact policy, became a rallying point for conservatives. Antitrust law had likewise reached its peak era of enforcement, with Justice Stewart lamenting, "the sole consistency that I can find is that in litigation under s 7 [of the Clayton Act], the Government always wins." Antitrust law provided Bork an opportunity to hitch his ideas to the battle for the courts being waged in the culture war. Just as a conservative backlash against Supreme Court "overreach" produced remarkable political success in the 1970s and 80s, so did Bork's critique of antitrust's vigorous enforcement.

The ammunition for Bork's offensive was readily available. For many, antitrust enforcement often meant "'coonskin cap' law

enforcement– the blind firing of muskets at companies that just seemed bad." In one incident, the Justice Department blocked a merger between two grocery chains that would have led to a combined market share of 7.5 percent. Comparisons between the justice department and an out of control frontier sheriff led the Justice Department to try to find the analog to "modern" forms of policing. Mainstream schools of antitrust, especially the Harvard school which had led the field previously, were forced to adopt many of the Chicago school's tenets, as lawyers and judges sought the "appearance of rigor" promised by Bork's ideas. It is perhaps harsh to say that the Chicago school of antitrust provided judges "an easy way to deal with hard cases," but the doctrine of focusing solely on prices led to a substantial deterioration of antitrust action. If a merger could not be shown to raise prices, for instance, it was allowed – regardless of whether it substantially reduced competition in the market in the long run.

The first casualties were per se (absolute) bans on vertical restraints, which were the most difficult bans to justify. As a flurry of legal challenges to antitrust rules produced victory after victory, robust enforcement of antitrust laws declined. Meaningful anti-merger action, for example, declined substantially in the post-Bork years. Ultimately the cost of this movement was nothing less than the spirit of antitrust, with entirely theoretical, often questionable economics being used to justify an attack on judicial precedent. After all, the research produced by the Chicago school was based purely in price theory, with little supporting empirical evidence, and centered on the assumption that "the existing structure is the most efficient structure." This assumption was as absurd as it was damaging for it assumed that anticompetitive effects "which did not exist in theory did not exist in practice." In other words, acts that had heretofore been considered anticompetitive must, if undertaken by a profit-maximizing firm, have been done to improve efficiency. This efficiency is (somehow) bound to result in lower prices for the consumer; therefore, what is in the interest of the would-be monopolist must (somehow) also be in the consumer interest; this was (somehow) obviously the intent of the framers of the antitrust laws. Such an outlook includes "preferences for economic models over facts, the tendency to assume that the free market mechanisms will cure all market imperfections, the belief that only efficiency matters," and finding a way to justify any non-intervention position. Taken together this attitude flies in the face of the main principles that had hitherto been enshrined in the Sherman Act, and that had been reiterated with the Clayton and Celler-Kefauver Acts. Nonetheless, its siren song of respectability and 'economic tools' to address difficult legal questions succeeded in swaying much of the legal establishment. All this came to a head during the 1970s, when antitrust lawyers and economists convinced the courts to "adopt an exclusively economic approach to antitrust questions" on an unprecedented level.

This reframing of the laws did not universally result in court decisions that lowered prices, which from the focus on price theory and "consumer welfare" one might think they would. Take resale price maintenance, a type of vertical restraint that manufacturers use to force retailers to sell at a specified (typically inflated) price, which was per se illegal until the 2007 Supreme Court decision in *Leegin vs. PSKS*. Advocates for resale price maintenance used economic theory to argue it provided consumers with better utility, and therefore better value. This was despite empirical evidence showing that states that at the time allowed resale price maintenance saw prices rise by 19% to

27%. It also ignored the potential for resale price maintenance to be used to disguise cartels. Even advocates of a laser-like focus on price levels could look the other way on occasion, preferring classical price theory to evidence. The Supreme Court ended up ruling that resale price maintenance ought to be judged on a rule-of-reason basis, which considers the procompetitive and anticompetitive effects of such a restraint, as well as the firm's market power to use such a restraint to entrench itself in the market.

More broadly, the rule-of-reason analysis is now applied near-universally, and its concern on maximizing consumer welfare is almost always about price. When this comes at the expense of competition, this presents a problem. After all, less competition might mean a more streamlined market, through say less sales and marketing costs, but this ignores that antitrust laws are fundamentally about preserving economic freedom and deconcentrating economic power. Furthermore, applying the Chicagoan doctrine precludes investigation of issues like income distribution, impacts on small businesses, and respect for civil rights, that all relate to how business is conducted. Indeed many of the anticompetitive effects of monopolistic practices that has previously been dismissed only saw robust academic validation in the post-Chicago era, with barriers to entry explored by Thomas Krattenmaker and Steven Salop, the use of patents to slow new market entrants shown by Carl Shapiro, the deleterious effects of mergers on innovation proven by Michael Katz and Howard Shelanski, and improved econometric models to answer theory with empirical evidence produced by Daniel Rubinfeld and others. Unfortunately it was too late, as antitrust enforcement entered an all-time low. The U.S. fought the last of its major dissolution cases against AT&T in the 1970s, and thereafter stopped pursuing major action. During the Reagan administration, the Justice Department aggressively pursued cartels but virtually stopped other forms of enforcement. An investigation by President Clinton's Justice Department into Microsoft, which controlled 90% of the operating system market and tried to corner the nascent internet browser market, became another casualty of the far-reaching election of 2000.

During the Bush administration, a record zero anti-monopoly antitrust cases were pursued, and no big mergers were blocked. Justice Antonin Scalia was arguing that "the mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free market system." Monopolies – in this case Verizon, which had held a century-long monopoly – were no longer evil, they were the unsung heroes of capitalism. And although Barack Obama promised "an antitrust division... that actually believes in antitrust law," the judiciary he faced had accepted many of Bork's ideas. Three quarters of American industries saw increased concentration between 1997 and 2012, 95% of mergers were allowed without further investigation. A swift-moving digital technology industry that in the 1990s promised compactness and an unprecedented level of creative destruction has consolidated into just a handful of sprawling giants. Facebook, an eight-year-old (old in the context of the internet) dominant firm, was allowed to buy one of its only real challengers, eighteen-month-old Instagram. There was no challenge from regulators on either side of the Atlantic, who absurdly concluded that the two were not competitors. Next came their purchase of WhatsApp, along with 65 more unchallenged acquisitions. Not to be outdone, Amazon and Google undertook 91 and 214 (of which only a couple came with conditions) respectively. Telephones, banking, cable, airlines,

pharmaceuticals, ticket sales, agriculture, food, and a whole host of industries have consolidated into a handful of players. AT&T, which was the last major breakup in the U.S., has seen its constituent parts re-group. The need for a return to a tradition of more robust antitrust enforcement is dire, and without action, things can only get worse.

Where We Go From Here

The current problems stemming from a more anticompetitive economy are not new; they are the same that prompted the first rounds of antitrust laws. Therefore, we need to revive the spirit of antitrust. This means cultivating a clear understanding of the dangers monopolies pose, not just to consumers' wallets but also to a functioning democracy. I am not proposing we abandon microeconomic analysis or scientific methods of tackling antitrust cases altogether. But getting rid of political considerations altogether because they are harder to deal with than microeconomics was a mistake. Robert Pitofsky warned us in 1979, as the Chicago school reached its zenith: "an antitrust policy that failed to take political concerns into account would be unresponsive to the will of Congress and out of touch with the rough political consensus that has supported antitrust enforcement for almost a century." Squabbling over whether a merger will increase the price of a service by a few cents per month risks ignoring the larger effects that consolidation has on society. Drawn-out court battles over the price effects of "moats" put up by firms with even some monopoly power can risk forgetting factors like innovation and the often-sinister reasons they are put into place to begin with. And treating internet giants with "free" services under the same rules as traditional firms must also be reconsidered.

Considering it has been nearly seventy years since the passing of the last landmark antitrust law, Congress needs to pass legislation updating antitrust for the current era. This means setting stricter rules for mergers and acquisitions, guidelines for dealing with firms that take users data instead of charging prices, specific criteria for investigating persistent monopolies, and procedures for breaking them up. To avoid any question of intent, the law should clearly state its democratic goals: preventing the centralization of private power and reaffirming to sovereignty of the people over the trusts. This will help the courts avoid the mistake of reinterpreting new antitrust legislation in ways that weaken its use.

Second, the public must be won over, for no substantial action can happen without its support. A good place to start is moving conversations around antitrust out of academia and the government and back into the public consciousness. There is substantial opportunity to do so: already Americans are three times more likely to express confidence in small business than big business, and this gap has only widened since the Great Recession. The tech monopolies in particular have been subject to greater scrutiny in recent years, which is reflected in Google and Facebook's record-breaking spending on lobbying in 2018 after a year of scandals. Economic inequality and the declining power of labor, which have become rallying cries for political reform, should be tied to antitrust. Political parties need to be held accountable for their antitrust agendas, and there is also good news on this front. The 2016 Democratic Party platform promised greater enforcement in order to "prevent excessively consolidated economic and political power, which can be corrosive to a healthy democracy." Matt Stoller, policy director at the Open Markets Institute, credits Senator Elizabeth Warren with this change. He argues that until a 2016 pro-antitrust speech, "monopoly wasn't

really an issue on people's radars," and after pressure from Senators Warren and Bernie Sanders, the party included stronger language in its 2017 "Better Deal" blueprint. There are signs the party rank-and-file are following this movement even as it threatens their longtime alliance with the tech industry, which gave twice as much to democratic candidates than republican candidates in the midterm elections. There is therefore evidence that political antitrust is back on the menu. Finally, government officials need to be held accountable on antitrust. Public awareness and scrutiny can make the bureaucrats in regulatory agencies accountable to their actions, as well as limit the damage dealt by the revolving door between industry and those who police it. Judges up for appointment should have their views on antitrust, and especially their interpretation of its purpose, thoroughly examined just as their positions on abortion, gun control, and the constitution are routinely scrutinized.

Third, the U.S. must pursue bold action against the persistent monopolies and force the dissolution of firms in heavily consolidated industries. This is the most important tool for any meaningful antitrust action, just as Roosevelt's trust-busting was the natural conclusion to the conditions of the Gilded Age. There is no reason to believe that breakups should be off-limits except in dire cases; for most of antitrust's history they were the default remedy to persistent monopolies. It is likewise fiction to claim that breakups are impossible. Sure, there are political costs involved, but logistically breakups are much like the spin-offs and reorganizations that have become popular in business. And while a monopoly might at first resist a breakup, it can be as good for the company as it can be transformative for the industry: consider the case of Standard Oil, whose constituent parts doubled in value just one year after its dissolution, and had quintupled in value just a few years later. The diseconomies of scale that harms consumers also holds back the giants and the stock markets that value them. There is therefore much to be said for the big case tradition. At worst, big cases can keep monopolists in line with a "policeman at the elbow," as seen by AT&T's avoidance of anticompetitive behavior while it was under investigation in the 1970s. At their best, breakups can transform entire industries, give rise to previously suppressed ones, lower prices, and boost stock market performance – all while safeguarding democracy from the danger posed by extreme concentrations of wealth.

Conclusion

Solving the antitrust problem will not make the larger problems in society disappear, nor is this essay the perfect guide for what should be done. Instead there needs to be a robust debate about the economic order the public wants to live in, and the direction the U.S. ought to take in this century. What is at stake is more than just the prices consumers pay at the register – it is one of the cornerstones of American democracy, which needs to be restored and bolstered once again. Just as the history of antitrust shows what happens when extreme concentrations of wealth are not reckoned with, it also shows that change is possible. We should look to this history to bring back the spirit of antitrust and strengthen our democracy.

References

[1] Act of July 2, 1890 (Sherman Anti-Trust Act), July 2, 1890; Enrolled Acts and Resolutions of Congress, 1789-1992; General Records of the United States Government; Record Group 11; National Archives.

- [2] Baker, Jonathan B. "The Case for Antitrust Enforcement." *Journal of Economic Perspectives*—Volume 17, Number 4 (Fall 2003).
- [3] Breyer, Stephen. *Leegin v. PSKS*. No. 06-480. Supreme Ct. of the US June 28, 2007. *Cornell Law*, January 20, 2019.
- [4] Crane, Daniel A. "All I Really Need to Know about Antitrust I Learned in 1912." *Iowa L. Rev.* 100, no. 5 (2015): 2025-38.
- [5] Crane, Daniel A. "Antitrust and Democracy: A Case Study From German Fascism." *University of Michigan L. Rev.* no. 595 (April 2018).
- [7] Crane, Daniel A. "Chicago, Post-Chicago, and Neo-Chicago." Benjamin N. Cardozo School of Law. Working paper no. 259 (April 2009).
- [8] Dave, Paresh. "Google Spends Big on U.S. Lobbying amid Antitrust, Bias Battles." *Yahoo! News*, Reuters, 22 Jan. 2019.
- [9] Jacobe, Dennis. "Americans Three Times as Confident in Small vs. Big Business." *Gallup.com*, Gallup, Inc., 27 July 2010.
- [10] Jeffreys, Diarmuid. *Hell's Cartel: IG Farben and the Making of Hitler's War Machine*. Metropolitan Books, 2008.
- [11] McGaughey, Ewan. "Fascism-Lite in America (or the Social Ideal of Donald Trump)." (August 2017). *British Journal of American Legal Studies*, Forthcoming.
- [12] Nachbar, Thomas B. "The Antitrust Constitution." *University of Virginia School of Law, Public Law and Legal Theory Research Paper Series* 2013-09, 2013.
- [13] Pitofsky, Robert. "The Political Content of Antitrust." 127 *U. Pa. L. Rev.* 1051 (1979).
- [14] Romm, Tony. "Pressure Is Building among Some Democrats for Another Antitrust Probe of Google." *The Washington Post*, WP Company, 4 June 2018.
- [15] Rothman, Lily. "How American Inequality in the Gilded Age Compares to Today." *Time.com*, TIME USA, 5 Feb. 2018.
- [16] Sohn-Rethel, Alfred. *The Economy and Class Structure of German Fascism*. Sohn-Rethel, Martin (trans.). Free Association Books, London (1987).
- [17] Stein, Jeff. "Warren's 2020 Agenda: Break up Monopolies, Give Workers Control over Corporations, Fight Drug Companies." *The Washington Post*, WP Company, 31 Dec. 2018.
- [18] "The 2016 Democratic Party Platform." *Democrats.org*, DNC Services Corporation, 2016.
- [19] Wu, Tim. *The Curse of Bigness: Antitrust in the New Gilded Age*. Columbia Global Reports, 2018.

The Encyclopedia of Porn or Diderot's Pornhub

DOROTHEA MOSMAN

UNIVERSITY OF OREGON

Porn gets called a lot of things—"information" isn't one of them. According to all manner of moment and movement porn is art, porn is devilry, porn is just porn. Yet wherever and whatever it is, it seems always to exist, it seems always to exist in excess, and it seems always to be considered as a social problem. The genesis and growth in the digital era of online porn has heightened the extent to which the preceding claims are true while increasing cultural attention to the problem of porn, on which Evangelicals remain united and academics as fractured as ever. The study of online porn, despite its relative novelty, serves as rich battleground for feminist thought: on the one hand, as recently reiterated by digital humanities scholar Rebecca Inez Saunders, the anonymity and easy access associated with online porn may be "considered...as the definitive beginning in a teleological chain that ends with sexualized children, increased sexual violence towards women, culturally vitiating obscenity and the commodification of the female body" (Saunders 236). As a rule, Saunders and others in her infant field aim, conversely, to "recognise the cultural and academic value of online pornography" (236). This new approach—and it is mine as well—takes for granted, first, that porn is out there; second, that a lot of porn is out there; and third, that something ought to be done to analyze, define, and perhaps eventually transform the immense social role it already plays in our lives.

Pornography: The Primer

In the feminist dialogue to which both Saunders and the field of porn studies itself remain deeply tied, porn dichotomously represents either a terminal site of women's utmost exploitation or a crucial conduit for their "sexual self-expression" (Saunders 236). Porn's social role in the popular consciousness is seldom so extreme, except where it is imagined by some conservatives as a scourge upon society; online porn, for most people, is just there to be looked at, sometimes tittered-over or jerked-off-to, but never read and rarely understood.¹ Porn, then, is only information when it acts as the *de facto* sex talk, or the natural supplement thereto, for pubescent boys (and girls) curious about what sex is. This is a scenario so far from uncommon as to be considered by sociologists Sabina, Wolak, and Finkelhor as "a normative experience" (*The Nature and Dynamics of Internet Pornography Exposure for Youth* 1). Conducting a survey of approximately 600 undergraduates at the University of New Hampshire in the spring of 2006, Sabina et al. gathered that "ninety-three percent of boys and 62% of girls were exposed to online pornography during adolescence," that "most exposure began when youth were ages 14 to 17," and that "boys were more likely to be exposed at an earlier age, to see more images, to see more extreme images (e.g., rape, child pornography), and to view pornography more often, while girls reported more involuntary exposure" (1). In these "normative" instances, porn's social role as information is made clear and significant. Porn is teaching young people, but especially young men, about sex, about power, about women.

¹"Porn," for the purposes of this paper, refers specifically to that which is online and videographic.

With this attitude toward online porn's social and educative role, Rebecca Saunders directs her focus in "The Pornographic Paratexts of Pornhub" toward "the biggest porn portal on the Internet" (Chafin 1). Though any scholar attempting a comparative analysis of the Internet's some four million porn sites will encounter "obvious variation in layout, content, rhetoric, objectives, and so on," says Saunders, "numerous aspects of Pornhub have become well-known and widely used, their consumer success establishing some of their prominent features as the mark of a professional porn site" (237). These features constitute the primary text to which Saunders devotes her analysis, culminating in her conclusion that Pornhub's "maturing paratextual elements" ultimately work against any meaningful shaping or reshaping of porn's social and/or feminist function. The question which remains in the wake of Saunders' close reading is this: If porn is information and Pornhub an archive—albeit, as Saunders deems it, a decentralized, anti-educatory, artless, essentially confused archive—how, then, should porn be organized?

For the answer I look not to the digital future but to a papered past, to the seventeenth century and Denis Diderot's *Encyclopédie*. In what follows I will work through Pornhub's structural, organizational, and feminist failures as outlined by Saunders and propose in their stead an Encyclopedia of porn (or, Diderot's Pornhub), based on the strategies of categorization—namely, alphabetization, editorial cross-reference, and expert curatorship—put forward in his 1755 article on the *Encyclopédie*. In keeping with Diderot's civic-mindedness, this proposal frames porn as information and the encyclopedization of Pornhub as a social project which centers and empowers women, both those actively involved in porn production as well as those affected by

emphasis and brackets my own).

The Worst-Laid Plans

In her contribution to 2014's *Examining Paratextual Theory and its Applications in Digital Culture*, Saunders nods to recent commentary from feminists, digital humanities scholars, and activist porn stars alike, which "offer[s] alternative, nuanced, and often positive interpretations of online pornography" (236). While she insists on online porn's feminist potential, Saunders does not locate that potential in Pornhub as is, and in fact, deconstructs the site's paratextual failures to live up to the anti-misogynistic, sex-positive ideal imagined by the likes of Sasha Grey. Maintaining that mainstream cultural condemnations of pornography interfere with porn's feminist potential, Saunders further argues that these "monolithically negative definitions of pornography may derive not from the hardcore content itself, but from the way in which the films are framed online" (237). It is Pornhub's failure to properly frame its porn films, regardless of their status as "hardcore" or otherwise, with which Saunders primarily takes issue.

Perhaps the most pressing problem with Pornhub's visual and philosophic organization is that it is visually and philosophically indiscriminate. Eighteen porn production companies supply the bulk of Pornhub's "seemingly infinite cache" of films, allowing Pornhub to host five-to-fifteen-minute excerpts from their full-length professional films in exchange for loose advertising rights (Saunders 237). Aiming to establish brand familiarity and lure viewers to their pay sites, these and other companies turn Pornhub into Adhub, where even in the films themselves block-lettered brand logos underline bobbing bodies. Rather than seek to limit or control the types of advertisements it hosts, Pornhub welcomes cross-promotion relatively undiscerningly. In a press release from 2014 spokespeople declared the site had "helped to boost the branding and exposure of content partners from all niches" transforming "traffic into monthly earnings!" (Saunders 239).

What might be called Pornhub's *laissez-faire* attitude toward advertisement is evident in its visual composition. While longtime users might build up preferences which narrow the types of ads they see to their specific tastes, new and newish viewers encounter ads which may seem disjointed, misdirected, or jarring. A novice visitor to the Pornhub homepage will encounter ads "from all niches"—but especially hardcore ones—well before he has clicked anything to suggest his interest in them. Far from attempting to mitigate the fundamentally discordant and distracting arrangement of these ads, Pornhub seems rather to favor the chaos. "Pornhub users are being centrifugally drawn away from the central visual text of the particular, chosen film," writes Saunders of the mishmash: "Many of Pornhub's adverts are for pay sites which offer more extreme, more apparently thrilling films, a sample of which are played on looped videos at an accelerated pace" (Saunders 246). Whether by accident or design, Pornhub's indifference toward to whom it sells ad-space and where on the site it situates those ads inadvertently focalizes violent, potentially misogynistic porn.

Pornhub articulates this indiscriminateness not only in its advertisements, nor indeed in how it juxtaposes those advertisements against its videos and other paratextual elements, but in its paratextual elements themselves. Saunders describes the details of this structural confusion:

"The hyperlinks which line the top of the homepage encourage users

to browse the categories, become a member, visit a related site, chat to other community members or see a live show via webcam. The rows of enthusiastically titled thumbnails which make up the bulk of Pornhub pages also push the user to exchange their existing text for something better, giving an enticing preview of their content as the cursor passes over them. Thus, seeking satisfaction in a single, feature-length film is discouraged. Adverts, hyperlinks and unceasing rows of thumbnails urge the user to be constantly, restlessly searching for something more exciting." (246)

Saunders does not likely intend the words "enticing" or "exciting" here—"extreme" or "hardcore" elsewhere—to read interchangeably with "misogynistic." But the crux of her argument relies on the implication that she does, at least in an overarching sense. Her emphasis on gratuity, excess, porousness, and indistinguishability as primary features of Pornhub's paratextual makeup is articulated, ultimately, as a feminist criticism, for she concludes that "Pornhub's construction of its films as individually valueless, as uniform, and as texts to be rapidly consumed and disposed of, facilitates the deeply embedded, sociohistorical notion that filmic pornography is culturally worthless" (247). This culturally engrained notion of porn as "worthless" is one she has previously castigated as unproductive and consequently antifeminist. She fails, however, to provide any viable proscriptions either for how Pornhub ought to change its anarchic structure on economic, philosophical, or feminist grounds, or for how feminists ought to approach the organization of Internet porn in the future.

A Diderotian Dialysis

The heretofore untackled feminist project which lies dormant in the space between Pornhub's slipshod thumbnails seems, if anything, an obvious move. Chaos calls out for order; information demands to be made sense of; Pornhub is practically screaming, not in the ecstasy it promises but in the panic it produces, for a remodel. The attempt to organize disorganized information (in this case, porn) concomitantly bestows upon its endeavor a heavy burden and a remarkable power. Writing in 1984, cultural historian Robert Darnton noted that she "who attempt[s] to redraw the boundaries of the world of knowledge would be tampering with the taboo" ("Philosophers Trim the Tree of Knowledge," 193). The taboo for this proposal is a foregone conclusion. But the prospect of reimagining and enforcing its subjects and borders is no less foreboding, its potential for social change no less pressing, than it was for Denis Diderot and Jean-Baptiste d'Alembert when with the *Encyclopédie* nearly three centuries ago "they undid the old order of knowledge and drew new lines" (Darnton 193). They did not take their power lightly. For Diderot the *Encyclopédie* was inexorably a social project with an ultimately pedagogical thrust; I have previously noted Diderot's assertion that the aim of an encyclopedia is to collect, structure, and transmit for future generations the preexisting archive of knowledge, "in order... that our grandsons, as they become better educated, may at the same time become more virtuous and more happy" (Diderot 291). Our grandsons' virtues remain, somewhat counterintuitively, a principal focus in the encyclopedization of porn, which will corral them ever so gradually out of (major aspects of) misogyny and into a feminist future. Diderot and I therefore share a social objective in our respective projects of molding better citizens. He attends with his "enormous task" to "man's curiosity, his duties, his needs and his pleasures"; so do I, with perhaps a greater emphasis

on the pleasures (Diderot 291). In order theoretically to “fix,” to functionally encyclopedize Pornhub, I will thus transpose three of Diderot’s most important strategies of Encyclopedic categorization onto the pornographic archive, namely, alphabetization, editorial cross-reference, and expert curatorship.

The first of these strategies is also the simplest. In “Philosophers Trim the Tree of Knowledge: The Epistemological Strategy of the Encyclopédie,” Robert Darnton maps out the ideological impetus behind the “tree” or “globe”² of human knowledge upon which Diderot and d’Alembert eventually alighted, while maintaining the ultimate arbitrariness of any such system. Although entries in the Encyclopédie were organized alphabetically rather than by subject, Diderot and d’Alembert nonetheless worked from a revised model of the tree of knowledge which attempted, like the versions put forth previously by Bacon, Chambers, and Locke, to organize the supposed fields of human knowledge into headings and subheadings all deriving from whatever its authors deemed to be the most fundamental phenomena of knowing.

And yet to determine what constitutes knowledge’s most fundamental phenomena is an impossible task which announces again and again its own illogic. The Encyclopedists as well as Darnton were conscious of this fallacy. “The very attempt to impose a new order on the world made the Encyclopedists conscious of the arbitrariness in all ordering,” notes Darnton—“What one philosopher had joined another could undo” (195). Describing their hesitance to favor any one formulation of the tree over another, Diderot admits in the Encyclopédie’s Prospectus that “The difficulty was greatest where it involved the most arbitrariness,” before launching into a philosophical lamentation:

“But how could there not be arbitrariness? Nature presents us only with particular things, infinite in number and without firmly established divisions. Everything shades off into everything else by imperceptible nuances. And if... there should appear a few [objects] that seem to break through the surface and to dominate the rest... they merely owe this advantage to particular systems... that have nothing to do with... the true institutions of philosophy” (Diderot, via Darnton 195).

Diderot and d’Alembert confronted these questions without much apparent resolution. Their decision to organize the entries of the Encyclopédie according to what Darnton calls “the innocent order of the alphabet” allowed them to avoid the pitfalls of organization by subject, which would have exposed their inescapably arbitrary tree. Behind the guise of “innocent” alphabetization, the Encyclopedists were also able to perform more subversive countercultural work than would otherwise have been possible (the alphabet obscured, for example, their model’s privileging of man over God). In its perfect arbitrariness, alphabetization was for Diderot less arbitrary than any other organizational model.

It is worth noting that Pornhub already has categories, which are, held up against the Encyclopédie’s, relatively explicit. Assuming a user manages to slip past the web of a dozen or more unrelated ads and videos vying for his attention from the homepage

² “Mappemonde,” notes Darnton on page 194, “was a crucial metaphor in [Diderot and d’Alembert’s] description of their work.” The metaphor was upheld by their project’s very name: “The word encyclopedia... derived from the Greek term for circle, signifying ‘concatenation of the sciences.’ Figuratively, it expressed the notion of a world of knowledge, which the Encyclopedists could circumnavigate and map.”

alone, he might click on the “Categories” tab and see them ranked, by default, according to “Most Popular,” and then, with a second click, alphabetically. It remains unclear, however, in what ways these established genres represent the fundamental phenomena of porn, or which great porn philosopher long ago decided so.³ Nor does it appear to matter much which category the user selects: regardless, the accompanying ads and suggested videos lean toward what Saunders describes as “more exciting” or “more extreme” (i.e., “up” rather than “down” the hardcore/softcore scale). Hence alphabetization alone is not enough to properly encyclopedize Pornhub.

The second Diderotian strategy of categorization vital to this proposal—editorial cross-reference—was hailed by Diderot himself as “the most important part of our encyclopedic scheme” (Diderot, *The Encyclopedia*, 310). In keeping with the project’s overarching civic-mindedness, Diderot acknowledges the cross references’ power to “suggest common elements and analogous principles,” but more important in his estimation is their ability to “confront one theory with a contrary one” and to “attack, undermine and secretly overthrow certain ridiculous opinions which no one would dare to oppose openly” (Diderot 310). Cross reference was, in its editors’ eyes, the Encyclopédie’s most compelling gesture toward social progress. Like carefully arranged hyperlinks on a webpage, cross-references represented for Diderot the perfect “method of putting men on the right path,” the right way “to lead people, by a series of tacit deductions, to the most daring conclusions” (Diderot 311). These “daring conclusions” often shared an antiestablishment thrust, with Diderot fashioning irony and insinuation into weapons against the church, the crown, and the aristocracy. Diderot famously linked “Anthropophagy” with “Eucharist,” “Communion,” and “Altar” (Darnton 200) and redirected readers of an entry on how to farm sugar to one which sweepingly condemned the institution of slavery (Werth 1). These examples are overshadowed in popular consciousness by the infamous “Capuchon” jab (which Diderot goes so far as to explicate in his article Encyclopédie), deciphered here by philosophy historian Colas Duflo:

“...The entry on Capuchon... appears at the end of the eulogy of the Cordeliers in the article Cordelier. The reference is an allusion to a ridiculous quarrel concerning the shape of the monks’ cowl that had divided this monastic order. The intention of the allusion is to tear the previous eulogy to pieces, but without printing a single reproachable word.” (Duflo 128)

So long as the cross references were “carried out artistically according to a plan carefully conceived in advance,” Diderot believed they would “give the Encyclopedia... the power to change men’s common way of thinking” (Diderot, “The Encyclopedia,” 311). While he insists that this elaborate subversive web functions only “when the author is impartial,” what he means is precisely the opposite. Only

³ They are as follows: 60FPS, Amateur, Anal, Arab, Asian, Babe, Babysitter, BBW, Behind the Scenes, Big Ass, Big Dick, Big Tits, Bisexual Male, Blonde, Blowjob, Bondage, Brazilian, British, Brunette, Bukkake, Cartoon, Casting, Celebrity, College, Compilation, Cosplay, Creampie, Cuckold, Cumshot, Czech, Described Video, Double Penetration, Ebony, Euro, Exclusive, Feet, Fetish, Fisting, French, Funny, Gangbang, Gay, German, Handjob, Hardcore, HD Porn, Hentai, Indian, Interactive, Interracial, Italian, Japanese, Korean, Latina, Lesbian, Massage, Masturbation, Mature, MILF, Music, Old/Young, Orgy, Parody, Party, Pissing, Popular With Women, Pornstar, POV, Public, Pussy Licking, Reality, Red Head, Role Play, Romantic, Rough Sex, Russian, School, SFW, Small Tits, Smoking, Solo Female, Solo Male, Squirt, Striptease, Tattooed Women, Teen, Threesome, Toys, Transgender, Verified Amateurs, Verified Couples, Verified Models, Vintage, Virtual Reality, and Webcam.

enlightened individuals (preferably with a capital 'E') who shared Diderot's own specific scientific, anti-nationalist, humanist, anti-classist sentiments would make proper cross-referencers.

It is with this deliberate emphasis on a highly editorialized cross-referencing system that I suggest the tactic be implemented in the Encyclopedia of Porn, whose general structure I will now propose. Just as Diderot moved implicitly and subversively within a semi-conventional model of human knowledge (the tree), this project will work within the framework of Pornhub's already-established categories⁴ to foster, first, a semblance of unity among the chronically disjointed films, and second, a trail of hyperlinks which purposefully guide the viewer toward less misogynistic and more socially constructive content.

The homepage of Diderot's Pornhub, then, will feature a sparse white background with only the hundred or so alphabetically-ordered categories and a small, expository thumbnail for each one displayed. Clicking on a category will bring the viewer to a primer vid: the most representative, straightforward (i.e., vanilla), professionally-produced video embodiment of said category, selected by editorial staff. Beneath the primer vid⁵, between three and five links to suggested videos will appear, each with a simple, descriptive title and thumbnail. Apart from the primer vid, the suggested video links, and the main "Pornhub" banner with its toggle of essential tabs (Home, Language options, Login/Sign up, Search, and Categories), no other content will appear on the screen. Advertisers interested in buying space on any one of these pages would submit their ad to a review process run by editorial staff, which will determine whether the content would be relevant and constructive on the proposed page—if it is, the advertisers may buy the spot; if it isn't, editorial staff may offer an alternative page on whose placement the content would be relevant and constructive, or simply reject the ad altogether.

Rather than orient the viewer "up" the hardcore scale by inundating him with ads and links to the "more extreme," the three to five suggested videos (cross references) would instead link to some equally kinky, some slightly less kinky, and some slightly kinkier related content in a gradual and not necessarily linear path to the extreme, should the user choose to seek it.^{6,7} Cross references in some cases would even, like Diderot's, connect unconventional concepts and material, thereby creating an impression in the user's brain of potentially new, socially constructive, progressive ideas through porn. Current Pornhub users who select "Anal" from the list of categories are instantly bombarded with links to videos like "Smoking bitch with awesome big tits takes it in the ass," "Banging Beauties Ass Fucked Sluts Kelly Divine Alexis Ford," and "Anal Revenge" (Pornhub, "Anal," 2018). Most are, to borrow

⁴ Future projects might work to curb or even eliminate some of Pornhub's more controversial categories, like racial fetishes, the sexual and cultural politics of which will have to be dealt with by a better-equipped author in a different paper.

⁵ In actuality, editorial staff will be continuously curating a range of primer vids for each category, such that frequent, even daily users aren't forced to watch the same videos and follow the same trail of hyperlinks every time they visit the site. The dog who plays Beethoven is actually seven dogs; the primer vid is actually seventy primer vids.

Saunders' verbiage, "triumphantly" degrading toward women, and all are heterosexual.

What if these weren't the first titles available to viewers interested in anal? Instead, a suggested video underneath the initial Anal primer vid might link, if not to good old-fashioned gay porn, to content in which a woman pegs her male partner, or, for that matter, her female one. A suggested video underneath the Interracial primer vid might link, unlike Pornhub, to something other than historically problematic and enduringly popular porn featuring black men "debasing" white women—instead, the hyperlinks could, for example, highlight films which depict interracial couplings where neither participant is white; where one or both are biracial; where both are women. Diversifying the content of suggested videos will particularly allow editors a unique opportunity to incorporate queer and trans pornography into the mainstream—to vanillaify it—rather than relegating it to a disproportionately stigmatized, ostensibly "deviant" niche in the pornographic landscape.

On the whole, these cross references will help porn perform its ideal social function as imagined by Rebecca Saunders, to render users "able to more actively contribute their own sexual self-expressions" and to "assert the validity of marginal sexualities" (236). Thus, where Diderot seized upon irony in his cross references in order to tear down corrupt institutions of power, to "snatch off the masks from the faces of certain grave personages," cross references in the Encyclopedia of Porn on the contrary will lift up marginalized voices, cultivating

positive associations with the oppressed rather than negative associations with oppressors (Diderot 311).

One obvious question this proposal has not yet answered relates to the composition of the new Pornhub's so-christened "editorial staff"; this, too, has to do with the third encyclopedic strategy I've borrowed from Diderot: expert curatorship. I have already outlined a few of the proposed responsibilities of these expert editors, but in order to highlight the crucial role they play in Diderot's social project and in my feminist one, I will here define who they are and how they are important to Diderot, and subsequently to the Encyclopedia of Porn.

When it comes to defining words, processes, and things, Diderot emphasizes that the lexicographer or encyclopedist must necessarily consult experts in the fields to which those terms pertain. For Diderot these "experts" are not limited to men of letters—although they, too, are included—but rather the craftsmen, artisans, and laborers of the world who daily perform the rituals

⁶ This assumes, of course, that the user in question is in fact looking for extreme/hardcore porn. Early site visitors might seek it [hardcore porn] out, at least in part because previous iterations of Pornhub have normalized and centered it. Some might be initially frustrated by its relative inaccessibility here. My hope is that gradually, as the encyclopedic system replaces Pornhub's current one and hardcore porn is deemphasized and denormalized, younger generations of viewers might not naturally gravitate toward violent porn.

⁷ The starting level of kinkiness for each page, and to a lesser extent the kink level of the accompanying links, would depend on the selected category. For example, the primer vid for the "Lesbian" category will likely be substantially more vanilla than the primer vid for the "Bukkake" category.

“...when it comes to questions, which this ultimately feminist project will center and will ask, of what makes porn good for women and what makes it bad, it's women, not their male directors, who should answer.”

which make a thing or a process what it is. His emphasis throughout the Encyclopedia remains rooted in the notion that no one man can cultivate enough knowledge to undertake such an ambitious project on his own, that collaboration is always imperative: “And who will furnish an exact definition of the word congruent unless it be a geometrician?...of the word epic unless it be a man of letters? of the word exchange unless it be a merchant?...of the word gouge, unless it be a man well-versed in the manual arts?” (Diderot 294). Refusing to privilege the contributions of the intellectual over those of the working man, Diderot instead posits a mutually beneficial partnership for and through the Encyclopedia, whereby each can learn from the other. Using his protocol, men of letters would commission “memorandum” (rough sketches of what would eventually become the Encyclopedia’s entries) from tradesmen. Though these working men’s memorandum will not be stylishly written, Diderot insists that they will “contain an infinite number of things which the most intelligent of men would never have perceived unaided, would never even have suspected, and hence could never have asked about” (Diderot 318). Diderot’s objective both in the creation and distribution of the Encyclopedia focused on building bridges between these two types of experts, unifying them into a progressive republic, educating them, letting them educate each other. “This is a work,” he puts forward, “that cannot be completed except by a society of men of letters and skilled workmen ... all bound together solely by their zeal for the best interests of the human race and a feeling of mutual good will” (Diderot 298; emphasis my own).

Inverting Diderot’s linguistic default to the masculine, we can imagine, in a feminist context, the collaboration of two types of experts—adult film actresses like Nina Hartley and the feminist academics like Rebecca Inez Saunders who study their profession—in the genesis of an Encyclopedia of Porn. Women like feminist porn star and entrepreneur Ela Darling, whose identities refute that binary, are especially welcome applicants to the editorial staff at Diderot’s Pornhub. “This is a work,” Denise Diderot might say, “that cannot be completed except by a society of women of letters and skilled working women ... all bound together solely by their zeal for the best interests of womankind and a feeling of mutual good will.” Who can supply a precise definition of the difference between bukkake and cumshot porn if not the actress whose specialty is one, the other, or both? Who better to define what maketh the MILF than a MILF? If primer pages (consisting of both the primer videos and suggested further reading) are this project’s equivalent to the Encyclopedia’s entries, then porn actresses are the tradeswomen and feminist scholars the women of letters—but in the ongoing, visual, digital project of Diderot’s Pornhub, the division of labor between the two is blurred, all are feminists, all are editors. This is not to say that male directors, producers, and actors in the porn industry cannot or ought not to be consulted when it comes to deciding on primer vids, deciding, essentially, what constitutes the most faithful, clear-cut filmic representation of a category like Cuckhold and what separates it from the porn genres against which it bumps up (and grinds)—merely that when it comes to questions, which this ultimately feminist project will center and will ask, of what makes porn good for women and what makes it bad, it’s women, not their male directors, who should answer.

On The Astral Plane

What will strike many as a remarkably simple principle in 2018—that women should be consulted about things which primarily involve

women—was radical in 1755. Through deigning to document their livelihoods alongside information more commonly associated with high culture and the upper classes, and furthermore to consult them in the documentation, Diderot sought to lend a small measure of the same cultural dignity and respect to the life of the working man which had so long been the exclusive prerogative of the rich. Cultural historian Raymond Birn traces the Encyclopedia’s editorial shift in the 1777-79 quarto edition away from Diderot’s encyclopedic vision of an honorable (French) working class which was worthy of being written and read about, toward less socially progressive goals. With a focus in “Words and Pictures: Diderot’s Vision and Publishers’ Perceptions of Popular and Learned Culture in the Encyclopédie” on illustrations in the original Encyclopedia, Birn examines which were kept and which were done away with in the truncated quarto edition. Unsurprisingly, mostly of those images that remained preserved “the facts of learned [high-brow] culture” while portraits of working-class life, which “Diderot had held so dear,” disappeared (Birn 73-74). “By rejecting the artisan’s world as a sphere of knowledge worthy of understanding and sympathy,” levies Birn, “[the quarto’s publishers] adapted the Encyclopédie to more traditional social and cultural norms” (74). The quarto edition was thus stripped of a great deal of its cultural meaning, just as the Encyclopedia of Porn would be were it to be dislodged from feminism.

The social project which had been the crux of Diderot’s vision for the Encyclopédie was abandoned. Where Diderot strove in the folio, as the Encyclopedia of Porn will, to “demonstrate truths, expose errors” and “skillfully discredit prejudices,” the quarto’s publishers strove, contrarily, to sell copies of a book to people who might buy it (Diderot 309).⁸ Diderot’s critics claimed he oversimplified, even romanticized peasant life in the Encyclopedia’s illustrations, but Birn replies that “A more fruitful analytical approach may be one sensitive to Diderot’s recognition of a hierarchy of labor and at the same time his desire to blur it in the service of dignifying the most menial of crafts” (79). As example Birn puts forward two very different illustrations of weavers which appeared next to each other in the folio that bore Diderot’s vision. The first depicts “the aristocratic art of tapestry manufacture in the royal Gobelins factory”—in this image the room is clean, well-lit, high-ceilinged; each employee occupies his own individual workspace; the subjects’ collective posture is good and their aspects are uniformly serene; “Every human, every object, has a sense of place in this controlled, sterile, hierarchical, pre-industrial environment” (Birn 79). The second illustration, of the “poverty-wracked” weavers of “laundry baskets, bread paniers, and cheap storage bins,” situates itself in a “dank and windowless” room inside which a handful of downtrodden laborers go to work. At least two are “emaciated,” one hunched over an unfinished basket, the other, “crippled or physically exhausted,” struggling up the stairs and half out of sight. These latter weavers could not look more at odds with their patrician counterparts—and yet “beside these images of surface misery,” Birn contends, “is a contrapuntal theme betraying creativity, intelligence, and imagination.”

How, then, in this scene of abject suffering, does beauty manage to seep through the cracks in the walls like the one beam of light slanting in from upstairs? Not how one might expect: “Three adults—we have no way of distinguishing boss from

⁸ Ironically, the quarto was supposed to be more affordable than its longer predecessor and therefore more likely to be purchased by readers outside the upper classes. It didn’t work out that way.

employee—have temporarily abandoned their customary labor to cooperate in an activity suffused with animation and creative pleasure. They are sculpting a Roman centurion in a stylized pose, thereby imitating an art of high culture comprehensible to the ENCYCLOPEDIA's leisured readership... Turned artistes, they work with naïve enthusiasm, overwhelm the misery of their surroundings, and convey to the beholder their dignity, sensibility, and artistic passion in a way that—unfamiliar with and likely contemptuous of their lives—he can understand.” (Birn 80)

Here Birn locates an almost impossible note of dignity, grace, and hope in what otherwise might legitimately be condemned as poverty porn. While he defends this potentially optimistic sliver of an otherwise gloomy scene, Birn simultaneously takes issue with Barthes' dismissal of the Encyclopedia's depiction of “the world of work” as unrealistically tame. “[Barthes] fails to convey yet another of Diderot's intentions,” says Birn, “—to win from readers a sympathetic comprehension of the dangers of mechanical labor” (81). An illustration of a miner fleeing from a potential explosion, Birn contends, “his face etched with fear of sudden, violent death” suggests Diderot was conscious of the occupational hazards—among them even possible death—which poor men confronted every day. He asks his wealthy readers, I would argue, not only to attempt to sympathize with the miner on the page before them, but to occupy, however fleetingly, however shoddily, his perspective: that fear for one's life, that fear of death, that fear of dying, that fear.

Where do porn stars and their Encyclopedia fit into these scenes of gentility and poverty, joy and fear, art and death? I have so far made little mention of how we ought to integrate the mass of low-budget, poorly-lit, thoroughly unempowering amateur porn on the Internet into the matrix of Diderot's Pornhub; porn whose leading lady smiles too much or not at all; sad porn, bleak porn, the unredeeming stuff. What I am advocating, like Diderot, is a shift in perspective. The encyclopedization of porn will help facilitate that move toward a larger feminist space within the realm of online porn by allowing women the editorial freedom to elevate porn which represents them fairly over that which does not, diminishing the harmful effects of young men internalizing misogynistic pornformation, and eventually, through the Encyclopedia's growth, creating more and more opportunities for women in the industry to produce and promote their own content. The quickest path to positive representation of women, after all, is when women are empowered to represent themselves.

The Diderotian encyclopedization of porn, equipped with its own distinctly feminist objectives and slant, can for its own purposes and within its own corner of the Internet center pornography which is not actively violent toward women, can even foster creative career spaces for women in porn to produce films which advance their own perspectives, but it cannot fundamentally dislocate the pornographic gaze which remains fixed on women as objects. It is not enough merely to pay the object better wages, let the object prioritize projects where it is made to feel like marginally less of an object than on others, make the object comfortable. If pornography is ever to be unitarily rather than fragmentedly, sporadically, anecdotally feminist, only a radical switch in perspective will do the trick. Consider as an experiment the unsmiling girl on the hotel mattress in a sort-of stranger's directorial debut. Now we must unstick ourselves from inside her co-star's bird's-eye-view and settle instead into a glassy-eyed perusal of the radiator; feel with her however briefly, however imperfectly, the pain, the embarrassment,

the rush; share whatever mindscape she is willing, for a moment, to share with us; experience with her the old emotion she and Diderot's miner know so well, the reason he's running, the reason she's not, of what it feels like to get fucked.

References

- [1] Chafin, Chris. “How the Internet Changed Porn.” *Rolling Stone*. September 12, 2017. Accessed June 4, 2018.
- [2] Darnton, Robert. “Philosophers Trim the Tree of Knowledge: The Epistemological Strategy of the Encyclopédie.” *The Great Cat Massacre: and Other Episodes in French Cultural History*. New York: Vintage Books, 1984. pp. 191-211.
- [3] Diderot, Denis. “The Encyclopedia.” From *The Many Faces of Philosophy: Reflections from Plato to Arendt*. Ed. Amelie Oksenberg Rorty. Oxford University Press, 2004. pp. 236-244. First published in 1755.
- [4] Duflo, Colas. “Diderot and the publicizing of censorship.” *The Use of Censorship in the Enlightenment*. Ed. Mogens Lærke. Boston: Brill Publishers, 2009. pp. 121-136.
- [5] Sabina, Chiara, Janis Wolak and David Finkelhor. “The nature and dynamics of internet pornography exposure for youth.” *Cyberpsychology and Behavior*, Volume 11, No. 6. 2008. pp. 691-693.
- [6] Saunders, Rebecca Inez. “The Pornographic Paratexts of Pornhub.” *Examining Paratextual Theory and its Applications in Digital Culture*. Ed. Nadine Desrochers. IGI Global, 2014. pp. 235-251.
- [7] Werth, Christopher. “At 300, Encyclopedia Pioneer May Yet Get A Hero's Burial.” *NPR*. October 5, 2013. Accessed June 7, 2018.

Involuntary Hospitalization And Bias Against Marginalized Groups

HENRY T. LYNCH

NORTHWESTERN UNIVERSITY

Similar to groups traditionally thought of as marginalized, such as ethnic groups, non-binary people, and women, purportedly mentally ill people are subjected to structural oppression. Despite being more likely to be victims of violence than perpetrators, we tend to think of mental patients as violent deviants, similar to the way black boys are consistently misidentified as being older and overall in possession of superhuman or subhuman traits. Already marginalized groups are disproportionately marginalized further by mental health care stigma and predatory insurance-seeking by health care providers. The administrative discretion psychiatric and law enforcement professionals are given to deal with mental patients or people suspected of having mental problems is effectively a license to incarcerate anyone at any time with no due process and no uniformly applied repercussions in place to deter abuses of power, and people incarcerated by law enforcement officials often cannot afford an attorney. The result is a sometimes-predatory system in which predatory professionals mask their coercive collection of people's insurance money by saying they are helping. Both the language we use—"cuckoo," "not all there," etc.--to talk about mental illness and the current structure of mental health care contribute to this further marginalization of the already marginal.

Stigma in Mental Health

How many times have you called--or heard someone called--"crazy?" It is so commonplace it often becomes disconnected from the actual group it references. This population group, the purportedly mentally ill, consistently referenced in day-to-day life more than any other population group, is currently subject to involuntary hospitalization, at any point, without a trial. A citizen needs only to call emergency workers and all but nominal rights are stripped away from the allegedly mentally ill person. In addition to lack of sufficient judicial oversight in the way people are involuntarily hospitalized, the facts of who is involuntarily hospitalized point to a systemic bias against already marginalized groups; mental patients are overwhelmingly poor, unemployed, and on welfare, and nonwhite males are more likely to be involuntarily hospitalized than white males [1][2]. City governments go too far in giving police and emergency responders unrestrained administrative discretion to break into a home without a warrant to take a purportedly mentally ill person against their will to a hospital under the pretext of potential harm. When such lack of restraint of the power to involuntarily commit someone is coupled with bias against people who have been previously hospitalized, socio-economic status, and race, then a potentially dangerous situation becomes a grave one.

Law Enforcement Officers' Administrative Discretion

There are strong similarities between the problem of militarized police departments and the treatment of the allegedly mentally ill. Much like the problem of the militarization of the police, as seen in the rise of "overwhelming paramilitary force," mild domestic disturbances have the potential to result in mandatory hospital stays of at least a few days when police are involved [3]. As an article covering police reactions to protesters at the 2009 G-20 summit said, "note that no one needed to

have broken actual laws to get arrested. The potential to break a law was more than enough. That standard was essentially a license for the police to arrest anyone, anywhere in the city, at any time, for any reason" [3](p. 12-13). For purportedly mentally ill people this describes their day-to-day life. Hospitalization can happen to anyone, anywhere, at any time, without any physical evidence of a reason. Neither first responders nor diagnoses can predict future acts of violence. Therefore, we cannot rely on psychiatric evaluations as the basis for incarcerating people.

Due Process

In addition to militarized police, unfairly withholding due process of law has also served as a mechanism of discrimination against purportedly mentally ill people. "Due process" is meant to protect against unfair proceedings involving restrictions of liberty in criminal courts, yet due process is virtually ignored in mental health courts. The growing body of mental health court documents is not available to the public, as releasing those documents would violate doctor patient confidentiality, among other rights; however, in a Utah district court case called A.E. and R.R. v. Mitchell, the court found no right-to-refuse treatment to exist [13]. In criminal court proceedings, defendants are jailed prior to their court date only they are a risk to themselves or to others. Otherwise, defendants are free to go until their court date. Courts are only involved in involuntary hospitalizations when a hospital decides to petition a judge for involuntary commitment after the mandatory "observation period." In California the observation period is 72 hours and can last weeks in other states. During this observation period hospitals collect insurance money with impunity, under no obligation to let the allegedly mentally ill person leave. Incidentally, this incentive of insurance money has gone as far as to lead hospitals to hire bounty hunters to round up people to fill hospital beds [4]. Furthermore, while people are involuntarily hospitalized,

people to fill hospital beds [4]. Furthermore, while people are involuntarily hospitalized, hospitals will often pressure patients to sign in voluntarily under the threat of petitioning a judge for a longer-term commitment [1]. If city officials are to comply with the U.S. Constitution's clauses about the right of citizens to due process of law, then the allegedly mentally ill ought to stand trial in a court before they can be hospitalized involuntarily.

The Misuse of Psychiatry in Courts of Law

There are certainly people who need to be separated from society, but this practice should be reserved for people who are proven to be violent. Mental health courts currently take the advice of physicians on almost every occasion. In practice, this fact means physicians have authority to incarcerate anyone. There is no one to check the physician, as judges are not medically trained. Mental illnesses are not visible on brain scans or through lab tests. This lack of tangible evidence leaves open the possibility of psychiatrists incarcerating people for superficial reasons such as the way someone presents themselves, the color of their skin, or their socio-economic status. Examples of psychiatrists using their power to oppress marginalized groups are widespread. In one such example, a study shows African Americans are more than three times as likely as whites to receive a schizophrenia diagnosis, and drapetomania was a diagnosis given to slaves as an explanation for why they tried to flee captivity [5][6]. These examples show how psychiatry is subjective and should have no place in a court of law, except perhaps as one piece of evidence among many other factors, and even then their opinion should not be viewed as more special or important than any subjective opinion. On the other hand, if there were verifiable proof of wrongdoing—a written statement of intent to commit an act of violence or witnesses who heard someone's intent to commit an act of violence—as is the requirement in courts of law, then there would be a basis for conviction visible to psychiatrists and lay people alike, lessening the potential for power abuse.

Implicit Bias and Language

Involuntary commitment has become commonplace just like racial bias in society is commonplace. In order to overcome the problem of involuntary commitment, it must be opposed in the same way. Oppression must be routed out of ourselves for true revolutionary change to take place [7]. While the stigma of mental illness is present at all levels of society, from people locked in mental institutions to people lightly called “crazy,” the individual is where the battle must be won. At the tame end of the spectrum, we use the word “crazy” to write people off. Just like police have a subconscious tendency to think of black people as sub- and superhuman, so too do we tend to ignore the humanity of the purportedly mentally ill [8]. There is the perception that the purportedly mentally ill are dangerous, when in reality the purportedly mentally ill are far more likely to be the victims of violence than the perpetrator [9]. At the extreme end of the spectrum the purportedly mentally ill have their liberties taken away right in our midst, that is, they are involuntarily hospitalized.

At the individual level the language we use could be changed—instead of “wacko” or “psycho” we could refer to people

as “distressed” or “in crisis.” Words such as “wacko” and “psycho” dehumanized and thus open the doors to large-scale systemic abuse. Furthermore, we could stop referring to people as “mentally ill” altogether, since the term “mental illness” is metaphorically referring to behavior deemed undesirable to society and does not refer to an actual biological disease visible with a brain scan or lab test. If police understood this fact they might be less inclined to bring someone to a hospital when there is a domestic disturbance and instead provide counseling or enforce laws regarding unacceptable behavior as needed. These names for the purportedly mentally ill also lead to those with an intersecting marginal identity to be further marginalized. Society is rife with implicit bias, and police are especially prone to act out implicit bias in detrimental ways due to the administrative discretion their position allows them and the disproportionate concentration of people—relative to the general population—in law enforcement who value the maintenance of hierarchical group superiority in their

interactions with others, also known as Social Dominance Orientation (SDO) [8]. SDO combines with administrative discretion and bias against the purportedly mentally ill, people of lower socio-economic status, and racial minorities to result in involuntary hospitalization just because someone called authorities about someone purportedly mentally ill. Empathy in our language is a protection against unwittingly creating a system structured to abuse and infringe on human rights.

“Empathy in our language is a protection against unwittingly creating a system structured to abuse and infringe on human rights.”

Equipping Law Enforcement Officers to Deal with the Distressed

A solution to oppression at the structural level is having police trained to deal with domestic situations involving purportedly mentally ill people without using force to bring them to a hospital. Police officers are trained to seize control of a situation when they think they might be dealing with someone armed or behaving erratically, often through stern, shouted commands. Shouting at someone and threatening to use force are not constructive ways to de-escalate situations with people in crisis. One in four people killed in officer-involved shootings are purportedly mentally ill [10]. We are clearly in dire need of Crisis Intervention Teams to de-escalate situations with the purportedly mentally ill. However, only fifteen percent of law enforcement agencies have crisis intervention training [11]. De-escalating situations could open up situations to an alternative to forced hospitalization. From a de-escalated situation non-coercive assertive community treatment programs could be put in place, programs that do everything it takes to keep people in the community and living independently, including helping people with housing, finances, and everyday problems in living. Programs like assertive community treatment pay for themselves by keeping people out of hospitals [12].

Behavior exists that poses a problem to society, but a fourfold solution should be applied to address it. First of all, we need to have transparency as to what is acceptable behavior and what is not. Having decisions about what is deemed acceptable behavior by society's standards concentrated in a single decider—be it emergency responders or psychiatrists—leads to chaos in the system as different deciders have different opinions about what is acceptable. Secondly, we

need for the legal system to stop using psychiatry as its underground, unofficial arm. If there are no public agreements on norms about behavior, no one can be held accountable for their actions, which is how the current system of psychiatry exists. Psychiatrists judge behaviors and either rule them acceptable or unacceptable based on their not widely circulated diagnostic manual, condemning people via diagnoses but not really holding them accountable because their behavior is then described as a disease. We need to officially enact laws pertaining to what behavior is unacceptable, which is the third part of the solution. These laws could then be enforced by police officers instead of giving police officers unrestrained administrative discretion to take people to hospitals, eliminating the acting out of bias against already marginalized groups such as African Americans and Latino Americans. Fourthly, we need to make these laws known. Legal education cannot be limited to lawyers if positive change is to occur.

Conclusion and Opportunities for Further Research

The misguided attempt of psychiatry and first responders to predict future harm, the disregard for the allegedly mentally ill's right to due process of law, and the subjectivity of psychiatry point to the necessity of abolishing the practice of incarceration for supposedly medical reasons. In order to end stigma at the structural level, that is, in courts, and at the level of day-to-day interactions, a revolution in thought must take place at the individual level and we must compassionately empathize with the allegedly mentally ill [7]. Sympathy leads to the medicalization—and further entrenchment—of oppression; empathy is needed to lead us out. Future research could investigate the use of counseling to de-escalate situations when police are called to deal with someone who is purportedly mentally ill. Non-coercive assertive community treatment could be examined as the alternative to involuntary hospitalization. Finally, there has been research about the positive effect descriptive representation has on minority racial groups, but there could be further research done into the effects of descriptive representation for the allegedly mentally ill.

References

- [1] Lurigio AJ, Lewis DA. Worlds That Fail: A Longitudinal Study of Urban Mental Patients. *Journal of Social Issues* 1989; 45: 79–90.
- [2] Rosenfield S. Race Differences in Involuntary Hospitalization: Psychiatric vs. Labeling Perspectives. *J Health Soc Behav* 1984; 25(1): 14-23.
- [3] Balko R. (2013, July 7). "Why did you shoot me? I was reading a book": The new warrior cop is out of control. Retrieved November 29, 2016, from http://www.salon.com/2013/07/07/why_did_you_shoot_me_i_was_reading_a_book_the_new_warrior_cop_is_out_of_control/
- [4] Payer L. Chapter 16/Psychiatric "Disorders" In: *Disease-Mongers: How Doctors, Drug Companies, and Insurers Are Making You Feel Sick*. NY: John Wiley & Sons, 1992: 234-235.
- [5] Schwartz RC, Blankenship DM. Racial disparities in psychotic disorder diagnosis: A review of empirical literature. *World Journal of Psychiatry* 2014; 4(4), 133–140.
- [6] Caplan AL, McCartney JJ, Sisti DA et al. *Health, Disease, and Illness: Concepts in Medicine*. Washington, D.C.: Georgetown University Press, 2004: back cover.
- [7] Collins PH. (2012). *Race, gender, sexuality, and social class: Dimensions of inequality*. Thousand Oaks, CA: SAGE.
- [8] Hall AV, Hall EV, Perry JL. *Black and blue: Exploring racial bias*

and law enforcement in the killings of unarmed black male civilians. *American Psychologist* 2016; 71(3):175-186.

- [9] Desmarais SL, Dorn RA, Johnson KL et al. Community Violence Perpetration and Victimization Among Adults With Mental Illnesses. *Am J Public Health* 2014; 104(12): 2342-2349.
- [10] Lowery W, Kinda K, Alexander KL et al. Distraught People, Deadly Results. *The Washington Post* 2015, June 30. Retrieved December 9, 2016, from <http://www.washingtonpost.com/sf/investigative/2015/06/30/distraught-people-deadly-results/>
- [11] Gleason L. (2014, April 29). NAMI Calls on Congress to Promote Nationwide Expansion of Police Crisis Intervention Teams (CIT). Retrieved December 10, 2016, from <https://www.nami.org/Press-Media/Press-Releases/2014/NAMI-Calls-on-Congress-to-Promote-Nationwide-Expan>
- [12] Latimer E. Economic considerations associated with assertive community treatment and supported employment for people with severe mental illness. *Journal of Psychiatry and Neuroscience* 2015; 30(5): 355–359.
- [13] Lebegue B, Clark LD. "Incompetence to Refuse Treatment: a Necessary Condition for Civil Commitment." *Am J Psychiatry* 1981, Abstract from www.ncbi.nlm.nih.gov/pubmed/7258384.

The Politics of Bilingualism in the United States: A New Perspective on the Immigration Debate

SARAH VICOL

STANFORD UNIVERSITY

So much of the rhetoric concerning immigrants today is not actually new. Once one takes a more expansive view of immigration to the United States, one begins to realize that immigrant stereotypes cross boundaries of race, ethnicity, and religion. But how can this be? The answer may lie in perhaps the most obvious yet somehow overlooked parallel between various immigrant groups. Many individuals in the past have arrived to the U.S. speaking very little to no English and subsequently do not give their native languages. The principal source of anti-immigrant sentiment might thus be rooted in language. Unfortunately, while there are plenty of separate studies of immigration and of bilingualism, there is a clear lack of scholarship that investigates the possible link between the two. This paper, while not highly rigorous, attempts to start filling in that gap. It will be shown that there exists a negative correlation between American attitudes towards bilingualism and immigration levels, and that this correlation boils down to three ranges of the foreign-born share of the U.S. population. It will also be shown that race and naturalization status are not confounding variables.

Introduction

It is an understatement to suggest that the 2016 U.S. presidential election was among the most fraught in recent history. It was clear from the moment that the eventual president announced his candidacy that immigration would be on the forefront. Indeed, immigration—or rather, the fear of it—might have very well determined the outcome of the election. President Trump, as both candidate and Commander-in-Chief, has successfully capitalized on what appears to be a growing anti-immigrant sentiment among the American public.

But what precisely is the root cause of this contemporary negative attitude? President Trump's hyperbole concerning Muslims and his administration's subsequent actions point to a potential religious aversion. But that would not explain the vitriol directed towards Hispanics and Latinos, the majority of whom are Catholic. One might be tempted then to identify ethnicity and race as the latent motivations. However, this line of reasoning, too, crumbles upon taking a more expansive view of immigration to the United States. While it is true that current immigrants are by far from Latin America, Africa, and Asia [1], that has not always been the case. For the majority of its existence, the U.S. received European immigrants [2], [3]. Yet this white and predominantly Christian group of people was also subjected to the same disdain that in the present day is associated with Muslims and Latinos [2], [4]. Catholics were discriminated against because they were the "wrong" kind of Christian (i.e., not Protestant) [5]. Benjamin Franklin labeled "Spaniards, Italians, French, Russians and Swedes" as "swarthy [in] Complexion" [6]. Italians, Irishmen, and even Poles were not considered to be white at one point in time [2], [7]. And Germans, whom today are considered the gold standard in efficiency and in strong work ethic, were considered lazy during the colonial era [4]. How can so many distinct cultures have at one point or another been branded with identical stereotypes? This would suggest that it is not something inherent to each individual immigrant group, but rather

something inherent to immigrants as a collective whole. Perhaps the most obvious yet somehow overlooked parallel between immigrants throughout the U.S.'s history—that many arrive speaking very little to no English and subsequently do not give up their mother tongues—might be the principal source of anti-immigrant sentiment. In other words, anti-immigrant attitudes could be rooted in language.

This paper will investigate the existence of a negative correlation between Americans' views of bilingualism and immigration levels over time; because this endeavor is a study and not an experiment, we will not be able to suggest a direction of causality. An opinion model predicated upon various official and personal documents, scientific studies, and poll results will be constructed and then compared against publicly available immigration data culled from government sources, primarily the U.S. Census Bureau. This undertaking should not be regarded as a rigorous longitudinal study; rather, it is meant to be a cursory review of current research compiled in a new manner and to serve as a possible stepping stone for future scholarship.

What we will be able to glean, if only in a preliminary way, from the methodology described above is that there is an inverse relationship between Americans' opinions of bilingualism and immigration rates over time. We will also be able to show that this negative correlation holds even when we account for potential confounding variables like race and naturalization status. It is important to remember that language—and by extension, bilingualism—cannot be studied in full isolation; our attitudes towards language and bilingualism are intertwined with our feelings regarding race, class, immigration in general, and other factors. But to make an argument about this complicated entanglement is not the purpose of this paper. Rather, we are interested in what different conclusions we might be able to draw upon determining how our views of bilingualism, disentangled as best as we can from some of the previously mentioned factors, have historically tracked with immigration levels.

Official Views of Bilingualism

There appears to be a sort of cyclical tendency in the way bilingualism has been historically viewed by the U.S. government and by American intellectuals. In the colonial era, the government did not mind the existence of bilingualism. Bilingualism was not actively promoted per se, but nonetheless it was not discouraged. In fact, the Continental Congress published its documents in English, French, and German. When John Adams proposed that English be made the official language of the United States, the other Founding Fathers demurred, suggesting that such an imposition would violate the rights of the individual. To this day, the U.S. does not have an official language; rather, English remains its de facto one. [2]

This sort of historical background is what we will use to contextualize what is referred to throughout this paper as our model of the U.S.'s official opinion of bilingualism. We will depict said opinion model as well as the eventual public and integrated societal ones as timelines. (See Figure 1.) Each timeline extends from the year 1800 until the present day in anticipation of the availability of immigration data; i.e., accurate immigration data is virtually nonexistent for the colonial era. Every decade will have a fill color, with red indicating negative views of bilingualism during a particular decade; yellow, neutral views; and green, positive views. If no data can be found for a particular decade, then said decade will remain colored white.

The official attitude towards bilingualism began to sour during the 19th century. (See Figure 2.) The 1800s saw the rise of nationalism and the conception of the nation-state, in which a country is not merely defined by territorial boundaries, but also by cultural and ethnic ones. The Prussian philosopher and linguist Wilhelm von Humboldt popularized the view that a defining characteristic of a nation-state was the existence of a single language spoken by that state's inhabitants. According to Humboldt, the "spirit of a nation" was encapsulated by its language [8]. When these ideas made their way to the United States, bilingualism among the lower classes began to be seen as evidence of a refusal to assimilate into society. Meanwhile, knowledge of multiple languages was still expected of members of the upper classes.

Figure 2. The U.S. government viewed bilingualism in a negative light during the 19th century.

But the upper classes' penchant for multilingualism waned in the early 20th century as the government became more strongly opposed to it. (See Figure 3.) Theodore Roosevelt was quoted in 1907 as saying that "[w]e have room for but one language in this country, and that is the English language" [7]. The U.S. sought to "Americanize" incoming immigrants like never before, and this entailed promoting "English only" [2]. As put by Ellwood P.

Cubberley, then-dean of Stanford University's School of Education, the first step in "break[ing] up immigrant groups ... to assimilate ... these people as part of our American race" was to make them give up their non-English native languages [2].

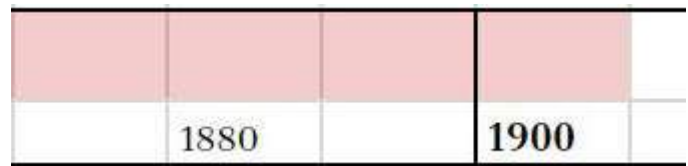


Figure 3. The first decade of the 20th century also saw the elite harbor negative opinions about bilingualism

World Wars I and II succeeded in embedding the belief into American society that even learning a foreign language was unpatriotic [2], [4], [7]. Nebraska banned the public use of languages other than English in 1919, and in 1923, Illinois declared that its official language was "American," though this was changed to English in 1969 [4]. A similar proposal barely failed in Congress in 1923 as well [9]. Figure 4 illustrates this continuing negative sentiment regarding bilingualism. Concurrently, a theory called linguistic relativity, also known as the Sapir-Whorf hypothesis or Whorfianism, was gaining credence within the academic world. Depending on how strictly this "hypothesis" was interpreted, it entailed that language either shapes or quite literally determines how one interacts with the world [10]. In other words, people who speak different languages experience the world differently from each other. This supposedly meant that one's intelligence was in part influenced by the language which one spoke; i.e., intelligence was inherently constrained by the "simplicity" of one's language. If a language could not express a certain idea, then how could an individual who only knew that language even

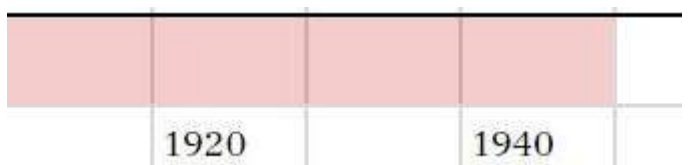


Figure 4. The U.S. government's negative attitudes towards bilingualism continued into the 1940s

understand, let alone devise, such a thought in the first place? Such notions meant that individuals could now claim superiority on the basis of language. Linguistic relativity and related theories



Figure 1. An empty timeline that will serve as the basis for the official opinion model

Figure 2. The U.S. government viewed bilingualism in a negative light during the 19th century

undoubtedly prejudiced early psychological research into bilingualism. Academics even claimed that bilingualism was a cause of mental retardation, arguing that the amount of brain “capacity” required to store two distinct languages systems meant less space for other knowledge and abilities. Studies purported to show that bilingual children were less intelligent than monolingual children [11], [12]. These results only compounded government officials’ convictions that it was in the U.S.’s best interests to suppress bilingualism, a sentiment only heightened by the Second Red Scare of the late 1940s and early 1950s. (Yes, even the specter of Communism manages to rear its head in this narrative.) In this context, bilingualism was framed as a possible national security threat. For example, a federal statute passed during this time not only delineated the trial and deportation procedures for illegal aliens and Communist “subversives,” but also amended the Nationality Act of 1940 to include English literacy as a requirement for naturalization

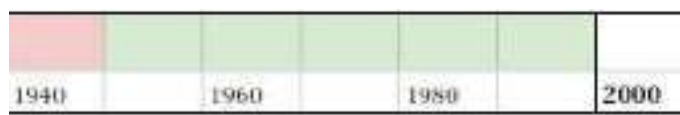


Figure 5. The net effect of the Cold War was positive official opinion of bilingualism for the rest of the 20th century

[13].

The latter half of the 20th century, however, saw a reversal in attitudes regarding bilingualism. (See Figure 5.) The U.S. government’s position on bilingualism evolved into a more favorable one as the country disavowed McCarthyism in the mid-1950s and searched for ways to beat the Soviet Union in the Space Race of the 1960s [14]. The fields of neuropsychology and neurolinguistics came into being, and researchers in these domains began to report that bilingualism came with a slew of benefits, including better executive function control and a delay in the onset of dementia [12], [15].

But once again the politics of bilingualism have been upended by recent events. Even though the number of Spanish speakers in America is decreasing due to intergenerational shift in language use among Hispanics—a phenomenon that occurs within any immigrant population [16]—there have been renewed calls for making English the official language of the U.S. [17]. New neuropsychological and neurolinguistic research suggests that previous reports of the benefits of bilingualism were overblown [12], [15]. And, in a modern-day though heavily diluted iteration of the Sapir-Whorf hypothesis, languages are now being “scientifically” categorized as “pleasant” or “ugly” [18]. As things stand now, the U.S. government and American intellectuals once more have negative

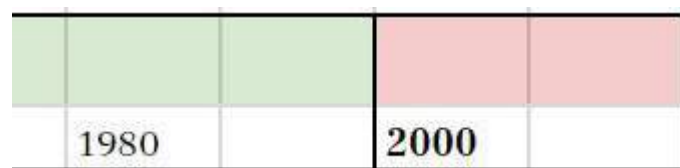


Figure 5. The final version of the official opinion model ends with the negative views of the last decade

views of bilingualism. (See Figure 6.)

Public Opinion of Bilingualism

While it is necessary that our model of the U.S.’s attitudes towards bilingualism over time be grounded in sufficient historical context, it would not be complete without insight into the opinions of the American public. Unfortunately, there is very little data appertaining to public attitudes towards bilingualism, so our model will only comprise the decades from the 1980s onwards.

To get a sense of the views of the general American public, we will first look at seven applicable polls conducted by the Pew Research Center. These polls, conducted somewhat infrequently beginning in April 1997, feature survey questions such as the ones listed below [19].

- Do immigrants have to speak English to say they are part of American society?
- How often do you personally come in contact with immigrants who speak little or no English—often, sometimes, rarely, never?
- Agree or disagree: It bothers me when I come in contact with immigrants who speak little or no English.
- Do most recent immigrants learn English within a reasonable amount of time?

We can amalgamate the responses to these various questions to formulate the public opinion model shown in Figure 7. The polling data suggests that public attitudes towards bilingualism were positive in the 1990s before declining to their nadir in the mid-2000s. Public sentiment improved around 2010 before again declining in 2015—when now-President Trump declared his candidacy, not so

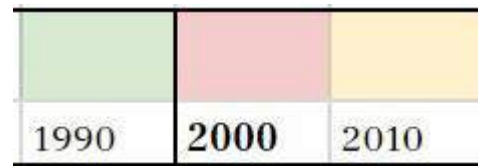


Figure 7. Public opinion of bilingualism fell from the 1990s to the 2000s. There is an overall neutral sentiment in the 2010s due to a considerable amount of fluctuation from positive to negative views

coincidentally—but is now back on the rise. (See Appendix.)

To strengthen our general public opinion model, we can delve into Americans’ Internet searches. Using Google Trends, we can explore potential correlations between bilingualism and immigration by using search terms as surrogates for the public’s beliefs. Figures 8 and 9 are screenshots of the comparison between monthly searches of “English only” and “bilingualism” and between “bilingualism,” “immigration good,” and “immigration bad,” respectively. These line graphs can both be divided into two time periods as indicated by the black dotted lines. From 2004 to 2011, there is a strong correlation between searches for “English only” and “bilingualism” and a weak correlation between searches for “bilingualism” and “immigration bad.” (See Figures 8 and 9, left of the dotted line.) We can surmise this as representing a period of time in which Americans had negative opinions of bilingualism. On the other hand, there has been no correlation between searches for “English only” and “bilingualism” since 2011, and there is instead a strong correlation between searches for “bilingualism” and “immigration good.” (See Figures 8 and 9, right of the dotted line.) This indicates that the public’s perceptions of bilingualism have been more favorable since the start of the current decade. Consolidating the results of Google Trends analysis does not

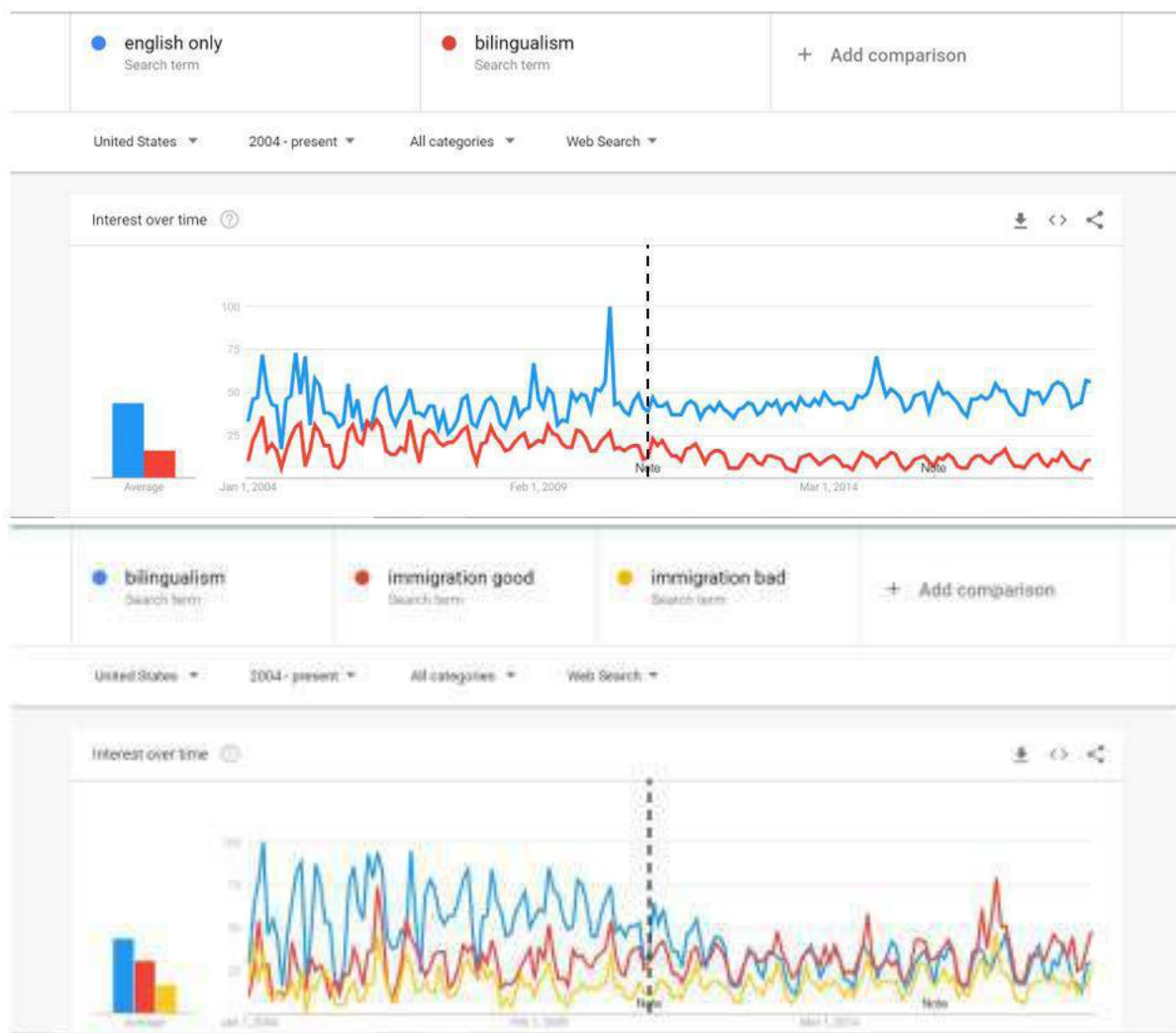


Figure 8. Comparison of monthly searches for "English only" versus "bilingualism" in the United States from January 2004 until the present day [20]

Figure 8. Comparison of monthly searches for "English only," "immigration good," and "immigration bad" in the United States from January 2004 until the present day [21]

alter our original public opinion model. (See Figure 10.)

We can broaden our public opinion model by considering the views of American bilinguals themselves. We will take the rate of intergenerational shift in language use among immigrants and their progeny as indicative of bilinguals' perception of their own bilingualism. Records from the 19th and 20th centuries led to the formulation of the three-generation model of intergenerational language shift; i.e., it was found that in the United States, knowledge of a foreign language is lost in about three generations [3]. While

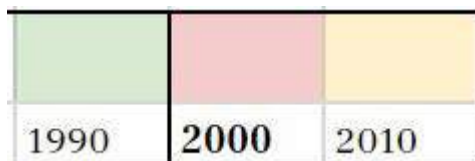


Figure 10. The public opinion model remains unaltered after adding Google Trends data

this is true in the present day [3], [16], [22], [23], it might not have always been the case. The three-generation model comes out of a convergence of existing data and is not an actual rigid rate. If society suddenly views bilingualism in a more negative light, then immigrants will encourage their children to speak English more often, resulting in knowledge of the native language dissipating in fewer generations.

Indeed, the three-generation model does not hold perfectly from the 1980s until about the start of the 2000s, but for the opposite reason. During these two decades, intergenerational language shift took longer than three generations, specifically amongst Hispanic- and Latino-Americans. That this occurred in the Hispanic population should be unsurprising, as Central and South Americans have constituted the largest group of immigrants since the 1980s [1]. In any case, this suggests that American bilinguals, and therefore the greater public, had positive attitudes towards bilingualism during the 1980s and 1990s. Incorporating this information into our public opinion model gives us the final iteration depicted in Figure 11.

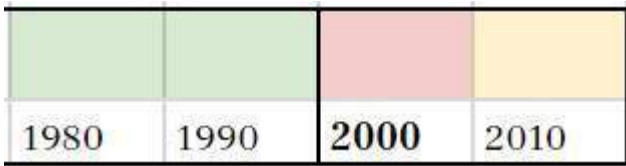


Figure 11. The finished public opinion model begins with the public's positive views of bilingualism in the 1980s

Comparison of Bilingualism Views and Immigration Levels

Before we can compare historical views of bilingualism against immigration data, we must integrate our official and public opinion models into a single societal model. (See Figures 12, 13, and 14, respectively.) It is apparent from seeing the official and public opinion timelines side-by-side that public opinion tends to follow the beliefs of the elite. This results in our integrated societal model appearing exactly the same as the model of the official view of bilingualism.

Figure 14. The societal model is the result of combining the official and public opinion models.

We will use the percentage of the total U.S. population that is foreign-born to represent what we interchangeably refer to as immigration levels or rates. We can overlay a line graph of this data with the corresponding portion of our societal model of bilingualism attitudes over time. (See Figure 15.) What we find is that there is indeed—and perhaps unsurprisingly—an inverse relationship (i.e., negative correlation) between views of bilingualism and immigration rates over time.

In fact, there is a far more definitive association that can be established from comparing our societal model against immigration levels. When the share of the total U.S. population that is foreign-born falls below 7.5 percent, Americans as a whole have a positive attitude towards bilingualism. When the foreign-born share rises above 10 percent, society has a negative outlook of bilingualism. The in-between range of 7.5 to 10 percent is ambiguous; the 1940s saw negative views of bilingualism although the proportion of the population that was foreign-born was less than 10 percent, and the 1990s saw positive opinion when said proportion was more than 7.5 percent.

With observational studies, we always run the risk of there being confounding variables. In other words, bilingualism and immigration may only seem to be correlated because there is some third factor that affects both of them in the same direction. One strong assumption that was made at the beginning of this paper



Figure 15. A line graph illustrating the foreign-born share of the U.S. population from 1850 until 2016 with corresponding bilingualism model colors overlaid by the author of this paper [24]

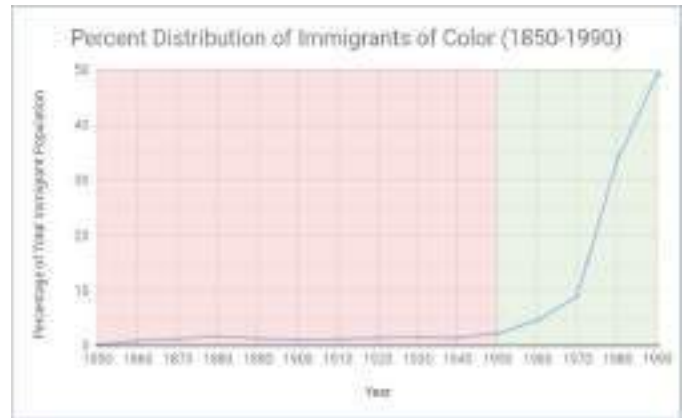


Figure 16. A line graph created by the author depicting the non-white share of immigrants from 1850 until 1990 and overlaid with corresponding bilingualism model colors [25]

involved the role of race in attitudes towards immigrants. We claimed that because the same stereotypes have been historically applied to many different racial and ethnic groups, race could not explain Americans' changing opinions about immigrants. Immigration data validates our reasoning. Figure 16 illustrates how the proportion of non-white immigrants has been increasing dramatically since



Figure 12. The completed official opinion model

Figure 13. The completed public opinion model

Figure 14. The societal model is the result of combining the official and public opinion models

the mid-20th century. If race were a confounding variable, then we would expect to see negative views of bilingualism during this time period and positive views otherwise. However, the very opposite phenomenon occurs. Therefore, our assumption that race is not a factor holds.

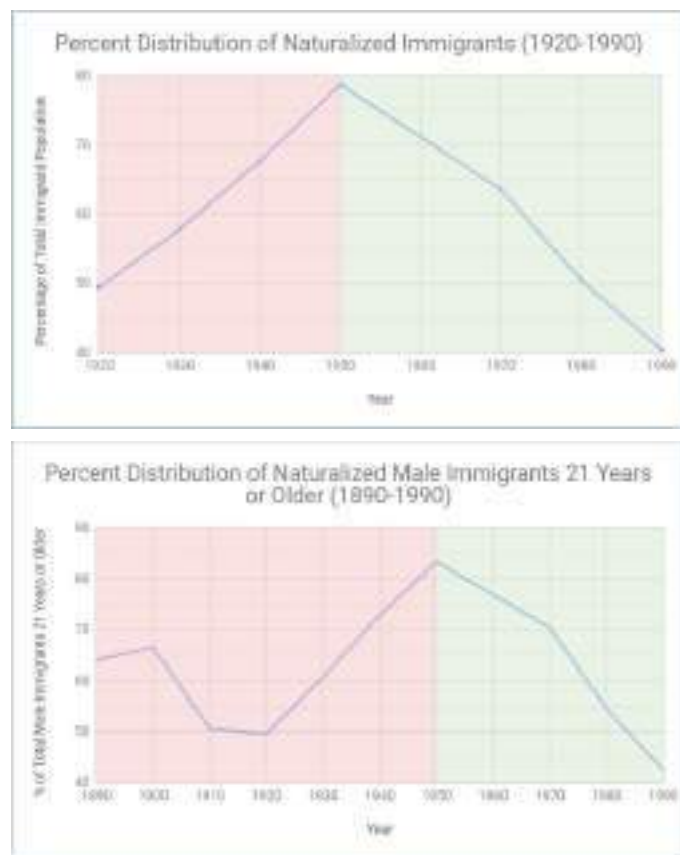


Figure 17. A line graph created by the author illustrating the naturalized share of immigrants from 1920 until 1900 and overlaid with corresponding bilingualism model colors [26]

Figure V. A line graph created by the author illustrating depicting the naturalized share of male immigrants 21 years or older from 1890 until 1990 and overlaid with corresponding bilingualism model colors [26]

Of course, this is not to say that language cannot have become racialized in the past few decades or that it was never racialized in the first place. As stated in the introduction of this paper, language and bilingualism cannot be studied in full isolation. But again, when we try to disentangle race from bilingualism as much as possible given our methodology, we find that race in general has no effect.

Another confounding variable might be naturalization status, the logic being that a decrease in the share of immigrants who are naturalized would be taken as a sign of diminishing assimilation, thus causing native-born Americans to turn against the immigrant population as a whole. If this were the case, then we would expect to see negative attitudes towards bilingualism when the proportion of naturalized immigrants falls. As with race, we find no such relationship. (See Figure 17.) Even when we compare views of bilingualism against the proportion of naturalized male immigrants

21 years or older—an immigrant subgroup for which there is data over a longer period of time—there still appears to be no correlation. (See Figure 18.) Therefore, we can conclude that naturalization status is also not a confounding variable.

In sum, comparing our integrated societal model of bilingualism attitudes over time against immigration data demonstrates that 1) there are positive views of bilingualism when the share of population that is foreign-born falls below 7.5 percent, 2) opinions of bilingualism are negative when said share of the population rises above 10 percent, and 3) race and naturalization status are not confounding variables.

Further Study

Academic and government research may be replete with separate studies of immigration and of bilingualism, but there is a clear lack of scholarship that investigates the possible link between the two in shaping our views on what it means to be an American. If an opinion model as imprecise as the one formulated in this paper still allows for some unequivocal conclusions to be drawn about the negative correlation between attitudes towards bilingualism and immigration levels, then only imagine what could be discovered with more meticulously created and exhaustive models. An obvious extension of this paper would hence be to conduct more rigorous studies using the same approach detailed here. Such studies would ideally use more data sets than does the model constructed in this paper, particularly when it comes to the public opinion component. (As an aside, it is hoped that public opinion polls concerning bilingualism will be conducted with greater frequency in the future, as this would greatly assist with such lines of inquiry.)

There are other avenues for further research that might be instigated by the results of this paper. It is crucial that confounding variables continue to be scrutinized and eliminated in order for the inverse relationship between bilingualism views and immigration rates to be made more concrete. Specifically, social class needs to be examined as a possible factor at play. A shared characteristic among immigrants, aside from speaking a foreign language, is social class; historically, most individuals who have immigrated to the United States have been poor. But poverty has always afflicted a sizable portion of Americans, and arguably the typical attitude towards poor native-born Americans is less hateful than towards poor immigrants. However, the potential influence of social class should not be discounted; what the analysis in this paper might have actually shown was not a negative correlation between opinions of bilingualism and immigration levels, but rather between social class and immigration levels, with bilingualism serving as a proxy for social class. Since the U.S. government does not directly record the social class of immigrants, it is suggested that educational attainment or some other similar attribute be used as an analog to social class.

Conclusion

Ultimately, we are nearing a fundamental crossroads in our nation's history. With an ever-growing population of Latinos [1] and of immigrants in general [24], it is imperative to determine whether the current heightened levels of contempt and hostility are historical anomalies or simply parts of a larger pattern. The fact that even this paper's relatively lax analysis revealed a bona fide inverse relationship between our attitudes towards bilingualism and immigration rates over time is telling. We need to conduct more statistically rigorous longitudinal studies in order to cement the notion that the

aforementioned negative correlation does indeed exist. Furthermore, potential confounding variables, namely social class, ought to be explored.

Margaret Thatcher perhaps best summarized the main difference between the United States and other Western nations: “Europe was created by history. America was created by philosophy” [27]. One of America’s driving philosophies has been equality—if not equality, then at least the pursuit of it. But right now, there is an increasing number of people in this country that are treated differently, and it may very well be because these individuals sound different. If we could explicitly identify the root cause of our historical ambivalence towards immigrants and foreigners, then perhaps we could finally begin to rectify the actual problem rather than its side effects.

Acknowledgements

Much gratitude is afforded to Dr. Tesla Schaeffer of the Program in Writing and Rhetoric for her unfailing guidance during the research process and during the preparation of this manuscript.

References

- [1] Regions of Birth for Immigrants in the United States, 1960-Present [Chart]. Migration Policy Institute (MPI) Data Hub. Available: <https://www.migrationpolicy.org/programs/data-hub/charts/regions-immigrant-birth-1960-present>
- [2] Betts B. Bilingualism in the United States: A Historical Perspective. Turlock: California State University at Stanislaus, 2004:1–3. Available: <https://www.csustan.edu/sites/default/files/TeacherEd/FacultyStaff/betts/Handouts/PDFs/Biling.%20in%20U.S.-A%20Hist.pdf>
- [3] Ortman JM, Stevens G. Shift Happens, But When? Inter- and Intra-Generational Language Shift Among Hispanic Americans. In: Population Association of America: 2008 Annual Meeting Program. Princeton: Princeton University, 2008:2–4. Available: <http://paa2008.princeton.edu/papers/80685>
- [4] Baron D. ‘Do You Speak American?’. Official American, English Only. MacNeil/Lehrer Productions 2005. Available: <http://www.pbs.org/speak/seatosea/officialamerican/englishonly/>
- [5] Lind D. Why historians are fighting about “No Irish Need Apply” signs — and why it matters. Vox Media 2015. Available: <https://www.vox.com/2015/3/17/8227175/>
- [6] Yglesias M. Swarthy Germans. The Atlantic 2008. Available: <https://www.theatlantic.com/politics/archive/2008/02/swarthy-germans/48324/>
- [7] Rosa RD. English-Only Movement. In: Thompson S, ed. Encyclopedia of Diversity and Social Justice (Vol. I). Lanham, MD: Rowman & Littlefield, 2015:295–6.
- [8] Mueller-Vollmer K, Messling M. Wilhelm von Humboldt. In: Zalta EN, ed. The Stanford Encyclopedia of Philosophy. Stanford: Stanford University, 2017. Available: <https://plato.stanford.edu/entries/wilhelm-humboldt/>
- [9] Mencken HL. The American Language; An Inquiry into the Development of English in the United States (4th ed., Abridged). New York City: Alfred A. Knopf, 1985:92–3.
- [10] Davis E. Does the World Look Different in Different Languages? New York City: New York University, 2016:1. Available: <https://cs.nyu.edu/davise/papers/Whorf.pdf>
- [11] Darcy NT. The effect of bilingualism upon the measurement of the intelligence of children of preschool age. The Journal of Educational Psychology 1946; 37(1):21–44. Available: <http://psycnet.apa.org/fulltext/1946-02116-001.pdf>
- [12] Konnikova M. Is Bilingualism Really an Advantage? The New Yorker 2015. Available: <https://www.newyorker.com/science/maria-konnikova/bilingual-advantage-aging-brain>
- [13] McCarran Internal Security Act, 64 Stat. 987 (Public Law 81-831). 1950.
- [14] Garcia O. Bilingualism in the United States: Present Attitudes in the Light of Past Policies. In: Greenbaum S, ed. The English Language Today: Public Attitudes Toward the English Language. Oxford: Pergamon Press, 1984:147–156. Available: <https://ofeliagarciadotorg.files.wordpress.com/2011/02/bilingualism-in-the-us-present-attitudes-past-policies.pdf>
- [15] Yong E. The Bitter Fight over the Benefits of Bilingualism. The Atlantic 2016. Available: <https://www.theatlantic.com/science/archive/2016/02/the-battle-over-bilingualism/462114/>
- [16] Spanish use is steady or dropping in US despite high Latino immigration. The Conversation 2018. Available: <https://theconversation.com/spanish-use-is-steady-or-dropping-in-us-despite-high-latino-immigration-85357>
- [17] 87% Say English Should Be U.S. Official Language. Rasmussen Reports 2010. Available: http://www.rasmussenreports.com/public_content/politics/general_politics/may_2010/87_say_english_should_be_u_s_official_language
- [18] Jenkin M. What makes a language attractive – its sound, national identity or familiarity? Available: <https://www.theguardian.com/education/2014/jul/17/what-makes-a-language-attractive>
- [19] Pew Research Poll Database. Available: <http://www.pewresearch.org/question-search/?keyword=English&x=0&y=0>
- [20] Monthly searches for “English only” versus “bilingualism.” Google Trends 2018. Available: <https://trends.google.com/trends/explore?date=all&geo=US&q=english%20only,bilingualism>
- [21] Monthly searches for “bilingualism” versus “immigration good” versus “immigration bad.” Google Trends 2018. Available: <https://trends.google.com/trends/explore?date=all&geo=US&q=bilingualism,immigration%20good,immigration%20bad>
- [22] Baran D. America’s Bilingual Roots. Language Magazine 2018. Available: <https://www.languagemagazine.com/2018/08/07/americas-bilingual-roots/>
- [23] Birner B. FAQ: Bilingualism. Linguistic Society 2012. Available: <https://www.linguisticsociety.org/resource/faq-what-bilingualism>
- [24] U.S. Immigrant Population and Share over Time, 1850-Present [Chart]. Migration Policy Institute (MPI) Data Hub. Available: <https://www.migrationpolicy.org/programs/data-hub/charts/immigrant-population-over-time>
- [25] Gibson C, Lennon E. Race and Hispanic Origin of the Foreign-Born Population: 1850 to 1990 [Chart]. In: Historical Census Statistics on the Foreign-Born Population of the United States: 1850-2000. Suitland, MD: U.S. Bureau of the Census, 2011. Available: <https://www.census.gov/population/www/documentation/twps0029/tab09.html>
- [26] Gibson C, Lennon E. Citizenship Status of the Foreign-Born Population: 1890 to 1950 and 1970 to 1990 [Chart]. In: Historical Census Statistics on the Foreign-Born Population of the United States: 1850-2000. Suitland, MD: U.S. Bureau of the Census, 2011. Available: <https://www.census.gov/population/www/documentation/twps0029/tab11.html>
- [27] Thatcher M. Speech at Hoover Institution Lunch. 1991. Available: <https://www.margarethatcher.org/document/108264>

Appendix

The following screenshots are of pertinent questions and their responses from the seven aforementioned polls conducted by the Pew Research Center. These questions address immigrants' proficiency in and use of the English language. [19]

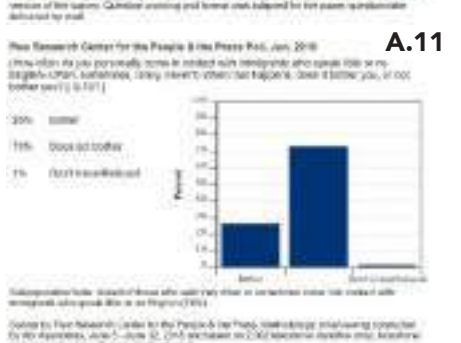
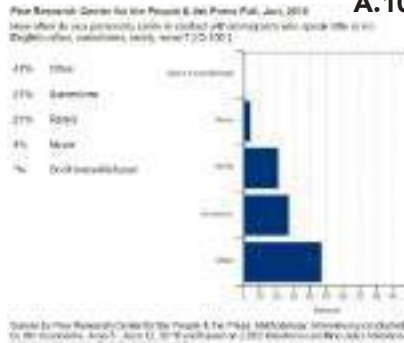
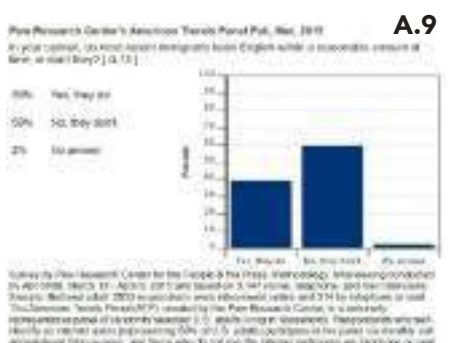
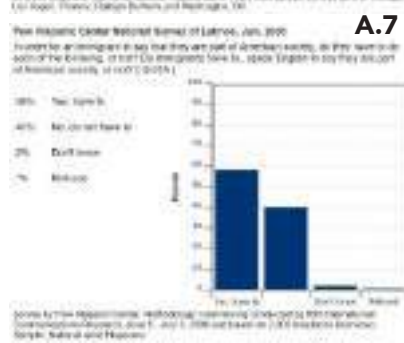
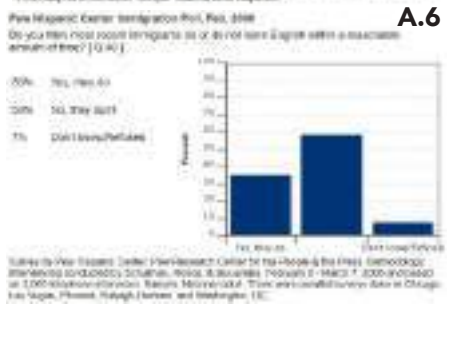
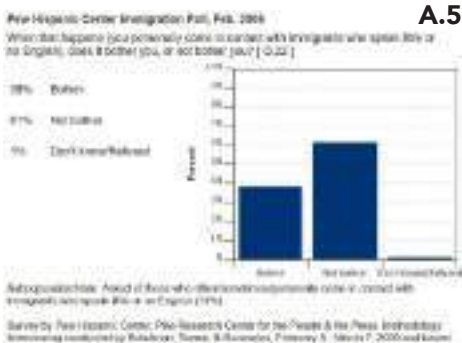
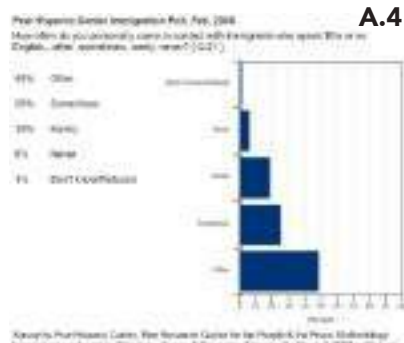
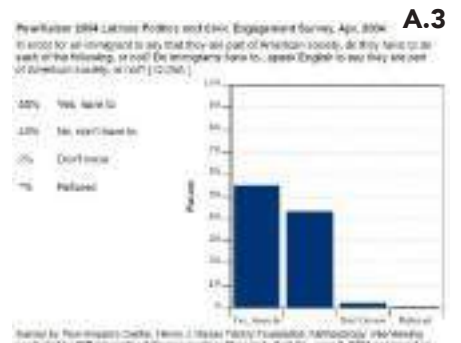
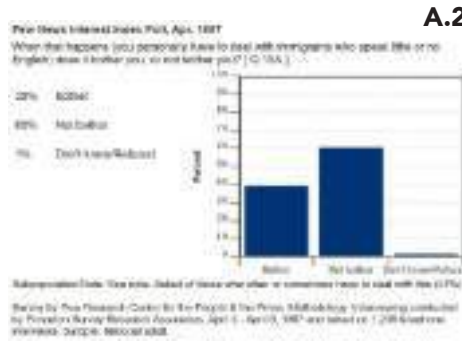
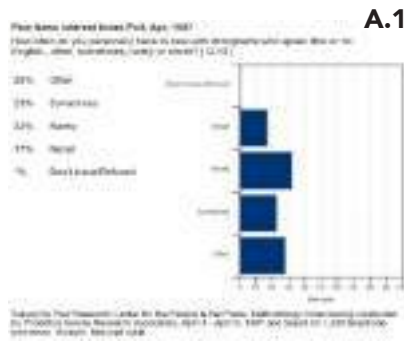


Figure A.1. Question 13 and its responses from an April 1997 poll
Figure A.2. Question 13A and its responses from an April 1997 poll
Figure A.3. Question 26 and its responses from an April 2004 poll
Figure A.4. Question 21 and its responses from a February 2006 poll
Figure A.5. Question 22 and its responses from a February 2006 poll
Figure A.6. Question 40 and its responses from a February 2006 poll
Figure A.7. Question 7A and its responses from a June 2006 poll
Figure A.8. Question 30Z and its responses from an April 2012 poll
Figure A.9. Question 13 and its responses from a March 2015 poll
Figure A.10. Question 100 and its responses from a June 2018 poll
Figure A.11. Question 101 and its responses from a June 2018 poll

Recurrent Control Nets as Central Pattern Generators for Deep Reinforcement Learning

VINCENT LIU, ADEMI ADENIJI, NATHANIEL LEE, JASON ZHAO
STANFORD UNIVERSITY

Central Pattern Generators (CPGs) are biological neural circuits capable of producing coordinated rhythmic outputs in the absence of rhythmic input. As a result, they are responsible for most rhythmic motion in living organisms. This rhythmic control is broadly applicable to fields such as locomotive robotics and medical devices. In this paper, we explore the possibility of creating a self-sustaining CPG network for reinforcement learning (RL) that learns rhythmic motion more efficiently and across more general environments than the current Multilayer Perceptron (MLP) baseline. Recent improvements in CPG modeling introduce the Structured Control Net (SCN), which maintains a standard MLP as the nonlinear module for global control but adds a linear module for local control [12]. SCNs are able to perform well on standard RL metrics, but struggle to produce coordinated locomotion as they are unable to capture time-dependent context. Here, we show that sequential architectures such as Recurrent Neural Networks (RNNs) model CPG-like behavior more effectively. Combining previous work with RNNs and SCNs, we introduce the Recurrent Control Net (RCN), which consists of a linear module for local control and an RNN as the nonlinear module for global control. We find that RCNs match and exceed the performance of baseline MLPs and SCNs across all environment tasks, confirming existing intuitions for RNNs on locomotive tasks and demonstrating the promise of SCN-like structures in control tasks.

Introduction

The ability to model rhythmic outputs given arrhythmic inputs has significant implications in neuroscience, robotics, and medicine. Central Pattern Generator (CPG) networks are a vastly important yet little understood region of reinforcement learning (RL). Existing baselines for locomotive tasks fail to explicitly provide rhythmic intuitions to RL agents.

In RL, the agent makes decisions based on a policy, which is a mapping from states to actions. During training, the model learns the mapping that will maximize reward. The standard policy network used when modeling locomotive tasks is the Multilayer Perceptron (MLP) [7, 9, 10]. OpenAI uses an MLP of two hidden layers and 64 hidden units as a baseline (MLP-64). This simple neural network learns to model tasks of moderate complexity, but fails to generate rhythmic output without rhythmic input [12].

While the MLP models global control relatively well, recent work has shown that explicitly modeling local control in addition to global control improves performance [5, 12]. [12] introduces the Structured Control Net, which outperforms the standard MLP across many environments. The SCN architecture is comprised of a linear module for local control and a nonlinear module for global control (in [12], an MLP-16), the outputs of which sum to produce the policy action. We use this baseline (SCN-16) for our benchmarks.

Previous work has explored the application of Recurrent Neural Networks (RNNs) to the central pattern generation task [15, 16]. RNNs maintain a hidden state that is updated at each timestep and allows RNNs to produce context-informed predictions based on previous inputs [6]. The vanilla RNN allows inputs and previous hidden states to flow between time-states freely and can provide more context to the action policy than the MLP and SCN can,

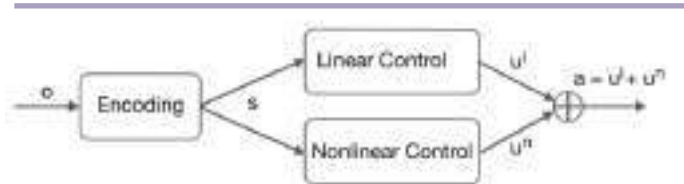


Figure 1. Structured Control Net architecture, adapted from [12].

since both only have access to information at the current timestep. This RNN architecture has many limitations, including loss of information over long time sequences, vanishing and exploding gradients, and complexity of parallelization [6]. We also explore the efficacy of variations to the vanilla RNN intended to mitigate these shortcomings, including Long Short-Term Memory (LSTM) [3, 8] and Gated Recurrent Units (GRU) [17], and provide results on the detrimental effect of increased RNN complexity on our RL environments.

In this paper, we combine the intuition behind SCNs and RNNs to exploit the advantages of modeling global and local control and of invoking global context at each timestep. We adopt the separation of linear and nonlinear modules from [12], which has been shown to improve performance by learning local and global interactions. We also adopt the vanilla RNN as our nonlinear module, which models global interactions more effectively than a MLP. We experimentally demonstrate that this architecture brings together the benefits of both linear, nonlinear, and recurrent policies by improving training sampling efficiency, final episodic reward, and generalization of learned policy, while learning to generate actions based on prior observations. We further validate our architecture with competitive results on simulations from OpenAI MuJoCo,

trained with the Evolution Strategies (ES) optimization algorithm [9, 13, 4].

Related Work

MLPs have previously been used to attempt modeling of rhythmic control tasks. The intuition is that the nonlinear fully-connected architecture is an effective function approximation. Although MLPs can generate high episodic rewards on many MuJoCo tasks, they often converge to locomotive behaviors that are jerky and unintuitive to motion.

Structured Control Nets

[12] demonstrates that enhancements can be made to the simple MLP model and yield boosts in performance across many environments. The SCN architecture learns local and global control separately. To model local interactions, the linear module is simply a linear mapping from the observation space to the action space. To model global interactions, the nonlinear module is a MLP, comprised of linear mappings as well as nonlinearities, giving it the ability to learn more complex interactions.

These interactions are specific to locomotion tasks: the agent needs to learn global patterns but also local interactions and movements specific to a task. Intuitively, the explicit modeling of local control is helpful because locomotive actions tend to depend on immediate prior actions. Although the SCN does not model cyclic actions, as it produces outputs given the current observations only, learning local interactions can provide more informative context than strictly learning global interactions. Therefore, we leverage the principle of separate control modules in our architecture

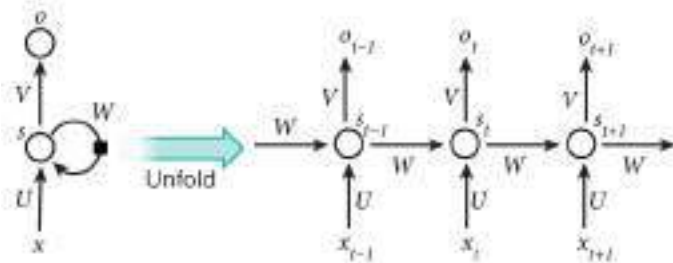


Figure 2. Recurrent Neural Network architecture

Recurrent Neural Networks

RNNs model time sequences well by maintaining a hidden state as a function of priors [6]. Traditionally, RNNs have been intensively used in natural language processing to model sequence prediction by including previous context. Therefore, to leverage the ability to condition current observations on past ones, RNNs have also been explored loosely in RL for quadruped environments. However, RNNs have not been generalized to general locomotive tasks and still remain relatively specific to quadruped-based locomotion tasks [16, 15]. While RNNs are inherently subject to a number of gradient and memory problems, there have been a number of modifications to the vanilla RNN architecture to address these problems. We explore some gated RNN variations in our experiments, but empirically opt for the vanilla architecture and will explain the intuition behind doing so later in this paper. (See Appendix A for a comprehensive description of effective RNN architectures.)

Experimental Setup

To work with locomotive tasks, we used OpenAI Gym [2], a simulated physics environment for RL, and ran our models on Multi-Joint dynamics with Contact (MuJoCo) tasks [14]. We used the MuJoCo ‘v2’ environments, which were the latest versions at the time of our experiments.

The Gym environment effectively serves as a wrapper to the MuJoCo tasks. At each timestep, the Gym environment returns an observation, which encodes the agent’s state (i.e. joint angles, joint velocities, and environment state). The policy takes this observation as input and outputs an action to be executed by the agent. In cyclic fashion, the environment returns the action, reward, and subsequent observation to the policy. Over many episodes and timesteps, the policy learns how to traverse the environment by maximizing the rewards of its actions.

Evolution Strategies

We used a population size per iteration of 20, a sigma noise of 0.1, and a learning rate of 0.01. We annealed the learning rate constantly throughout training with a decay factor of 0.999. We also used an epsilon-greedy policy with an initial random exploration rate of 1. We linearly annealed this probability to 0 over 1 million timesteps.

Recurrent Control Net

We built upon the concept of separate linear and nonlinear modules from [12] and designed our Recurrent Control Net (RCN) in a similar fashion. Our linear module is identical to that of the SCN [12], but our nonlinear module is a standard vanilla RNN with hidden size 32. Intuitively, the linear module provides local control while the nonlinear module provides global control. However, unlike the MLP used in SCN-16 [12], the RNN learns global control with access to prior information encoded in its hidden state. We used this architecture (RCN-32) as our baseline in experiments.

In our experiments, we compared the RCN-32 to a vanilla RNN of hidden size 32 (RNN-32) to test the efficacy of the extra linear module. To reduce the number of trainable parameters for ES, we removed all bias vectors in all models.

Evaluation

We used OpenAI’s MLP-64 model and the SCN-16 outlined in [12] as baselines for experimental comparison. We evaluated a model’s efficacy by its final reward after 10 million timesteps rather than the rate of convergence. Across all MuJoCo environments, we find that the RNN-32 matches or exceeds both baselines (see Figure 3). We also notice that the RCN-32 consistently improves upon the RNN-32. We only show the results across the Walker2d, Swimmer, Humanoid, and Half-Cheetah environments as they best represent locomotive tasks (as opposed to Humanoid-Standup and Reacher, for example) and provide the most interesting training curves across all models (as opposed to the Ant environment, in which all models converge to negative rewards).

From our experimentation with various recurrent structures, we make several interesting observations. The recurrent structure seems to be inherently conducive to modeling locomotive tasks because its hidden state explicitly encodes past observations, whereas an MLP does so implicitly. We desire this explicit encoding because it facilitates learning of patterns in sequential observations. We also found that the increase in model complexity past a certain threshold is detrimental to ES’s randomized training process. Additionally, explicit modeling of linear and global interactions

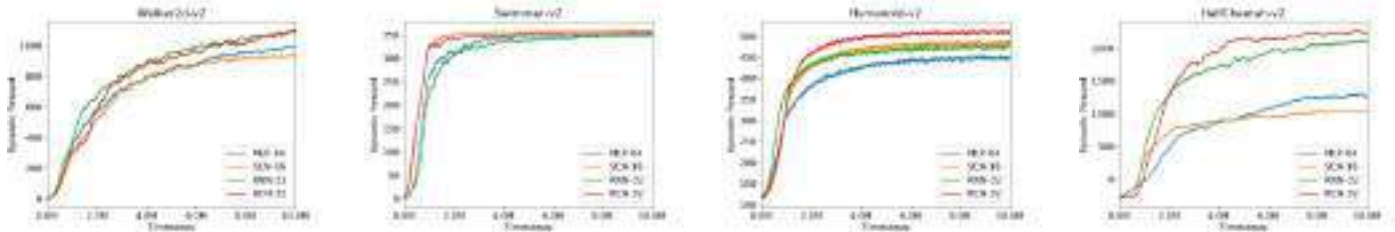


Figure 3. Episodic rewards for MuJoCo environments on baselines MLP-64, SCN-16, RNN-32, and RCN-32 using ES optimization. Average of 5 median trials from 10 total trials.

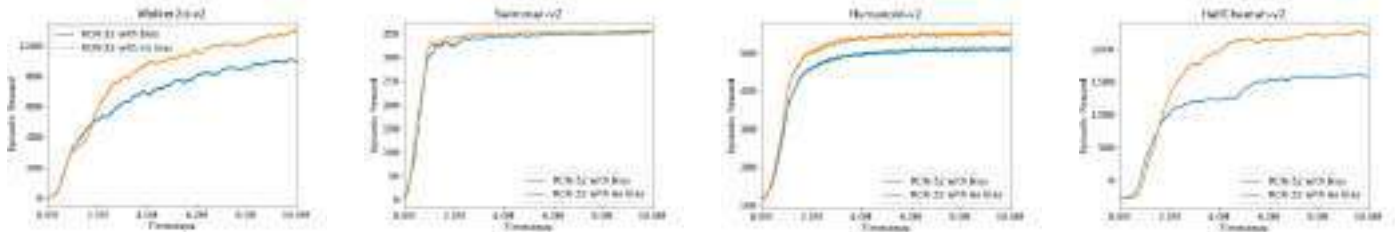


Figure 4. Episodic rewards on MuJoCo environments with RCN-32, with and without biases using ES optimization. Average of 5 median trials from 10 total trials

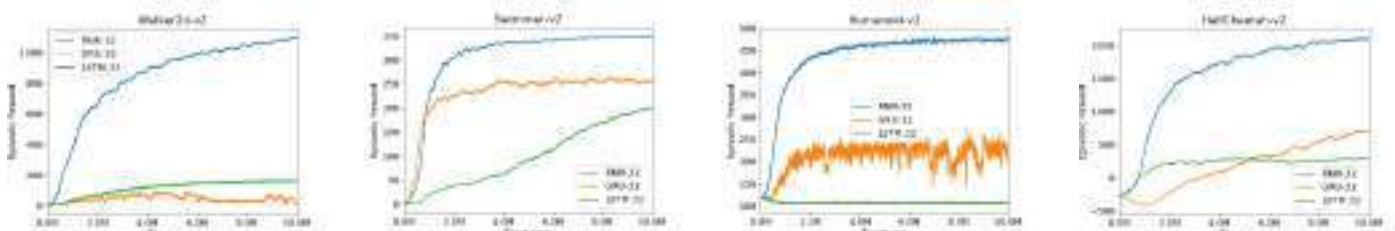


Figure 4. Episodic rewards on MuJoCo environments with RNN-32, GRU-32, and LSTM-32 using ES optimization. Average of 5 median trials from 10 total trials.

with linear and nonlinear modules consistently improves model performance.

Gated Information Flow

In all our trials with ES optimization, we noticed that recurrent architectures with gated information flow (GRUs, LSTMs) struggled in training (see Figure 5). We believe that since ES is a random optimizer, it struggles to optimize models with more parameters. A more complex model introduces more local optima, which may cause ES to converge prematurely. Additionally, since MuJoCo tasks are relatively simple and low-dimensional, enhanced memory is unnecessary and the learning of gates in training only burdens the optimization process.

The ES algorithm is inherently hampered by its gradient-free approach. Because it updates weights with random noise, models with more parameters are subject to higher overall noise variance per iteration. This can cause complex models to fail to converge entirely (see Figure 5). However, with simple architectures, we see early convergence in episodic reward (compared to the same models trained with different algorithms). As such, we anticipate that GRUs and LSTMs may achieve higher rewards with an optimization algorithm like PPO (Proximal Policy Optimization, a gradient-based optimizer [11]), where extra parameters from information gates are not heavily penalized.

Linear Control

The RCN-32 consistently outperformed the RNN-32. This finding is

consistent with [12], which shows the efficacy of introducing a linear component in addition to the nonlinear component (see Figure 3). This tiered approach accounts both for immediate information provided per observation and for longer-term patterns. As we have mentioned before, individual actions in locomotive tasks are heavily conditioned on immediate observations. The separate linear module allows for a larger emphasis on local information.

Local control is balanced by the RNN, which is responsible for global control. The hidden state is a complex series of nonlinear mappings of past observations, which gives the RNN access to global information. The addition of the linear module to the nonlinear module allows the entire architecture to learn local and global interactions. While the increase in performance is sometimes marginal, the SCN-16 similarly improved the MLP-64.

Incorporating Biases

We also experimented with incorporating biases into the RCN. Doing so immediately decreased performance across all tasks, sometimes even below baseline performances (see Figure 4). We believe that this is because the inclusion of biases burdened the optimizer in training without providing any real value to what the model learns. Just as the gated RNN variations struggled in training, adding parameters to the RCN resulted in higher noise variance per ES iteration. Another possible explanation is the simplicity of the MuJoCo environments, which may not require the additional bias vector to successfully model the task.

Conclusion

We conclude that RNNs model locomotive tasks effectively. Furthermore, we conclude that the separation of linear and nonlinear control modules improves performance. The RCN combines the benefits of both concepts, learning local and global control and patterns from prior sequential inputs. We also note the detriment of increasing model complexity with information gates, though this is probably due to MuJoCo task simplicity and the ES training algorithm. Because ES updates weights randomly, additional gates create more local optima that ES has to overcome.

Since our models have only been trained with the ES algorithm, future investigations would involve exploring the performance of RCNs with an algorithm such as PPO. Additionally, recent practices in natural language processing have successfully replacing recurrent layers with convolutional layers [1]. It would be interesting to explore whether convolution could replace the RNN module for sequential modeling. We hope that our findings open up further investigation into the usage of RCNs for these applications.

Appendix A: Recurrent Architectures

For more context, this section covers in-depth the fundamental recurrent architectures upon which we built our models: Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs) and Long Short-Term Memories (LSTMs).

Recurrent Neural Network

The vanilla RNN maintains an internal hidden state to compute future actions, which serves as a memory of past observations. This simple architecture allows all inputs and hidden states to flow freely between timesteps. Standard RNN update equations are below.

Gated Recurrent Unit

$$\begin{aligned} h^{(t)} &= \tanh(W_h h^{(t-1)} + W_x x^{(t)} + b_h) \\ o^{(t)} &= W_o h^{(t)} + b_o \end{aligned}$$

Equation 1: Vanilla Recurrent Neural Network update equations, where $h(t)$, $o(t)$, $x(t)$ denote the hidden state, output (action), and input (observation) vectors, respectively, at timestep t .

A GRU improves upon the vanilla RNN by learning to retain context for the next action by controlling the exchange of inputs and previous hidden states between timesteps [17]. GRUs have a reset gate r after the previous activation to forget part of the previous state and an update gate u decides how much of the next activation to use for updating.

Long Short-term Memory

An LSTM learns a “memory” of important locomotion context via input, forget, and output gates [3, 8]. The input gate i regulates how much of the new cell state to keep, the forget gate f regulates how much of the existing memory to forget, and the output gate o regulates how much of the cell state should be exposed to the next layers of the network.

Acknowledgements

$$\begin{aligned} u^{(t)} &= \sigma(W_{u,x}x^{(t)} + W_{u,h}h^{(t-1)} + b_u) \\ r^{(t)} &= \sigma(W_{r,x}x^{(t)} + W_{r,h}h^{(t-1)} + b_r) \\ h^{(t)} &= u^{(t)} \circ h^{(t-1)} + (\bar{1} - u^{(t)}) \\ &\quad \circ \tanh(W_{h,x}x^{(t)} + W_{h,h}h^{(t-1)} + b_h) \\ o^{(t)} &= W_{o,h}h^{(t)} + b_o \end{aligned}$$

Equation 2: Gated Recurrent Unit update equations, where $h(t)$, $c(t)$, and $x(t)$ denote the hidden state, cell state, and input (observation) vectors, respectively, at timestep t . The output is produced with a linear mapping of $h(t)$ to the output (action) vector.

$$\begin{aligned} f^{(t)} &= \sigma(W_{f,x}x^{(t)} + W_{f,h}h^{(t-1)} + b_f) \\ i^{(t)} &= \sigma(W_{i,x}x^{(t)} + W_{i,h}h^{(t-1)} + b_i) \\ o^{(t)} &= \sigma(W_{o,x}x^{(t)} + W_{o,h}h^{(t-1)} + b_o) \\ c^{(t)} &= f^{(t)} \circ c^{(t-1)} + i^{(t)} \\ &\quad \circ \tanh(W_{c,x}x^{(t)} + W_{c,h}h^{(t-1)} + b_c) \\ h^{(t)} &= o^{(t)} \circ \sigma(c^{(t)}) \end{aligned}$$

Equation 3: Long Short-Term Memory update equations, where $h(t)$, $c(t)$, and $x(t)$ denote the hidden state, cell state, and input (observation) vectors, respectively, at timestep t . The output is produced with a linear mapping of $h(t)$ to the output (action) vector.

This research on modeling central pattern generators originally began as a class project for CS 229, Machine Learning. We thank our TA advisor, Mario Srouji, for giving us insight on his Structured Control Net [12].

References

- [1] Eric Battenberg, Jitong Chen, Rewon Child, Adam Coates, Yashesh Gaur Yi Li, Hairong Liu, Sanjeev Satheesh, Anuroop Sriram, and Zhenyao Zhu. Exploring neural transducers for end-to-end speech recognition. In Automatic Speech Recognition and Understanding Workshop (ASRU), 2017 IEEE, pages 206–213. IEEE, 2017.
- [2] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. arXiv preprint arXiv:1606.01540, 2016.
- [3] Yang-Hun Cha, Dang-Yang Lee, and In-Kwan Lee. Path prediction using lstm network for redirected walking. In 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pages 527–528. IEEE, 2018.
- [4] Edoardo Conti, Vashisht Madhavan, Felipe Petroski Such, Joel Lehman, Kenneth Stanley, and Jeff Clune. Improving exploration in evolution strategies for deep reinforcement learning via a population of novelty-seeking agents. In Advances in Neural Information Processing Systems, pages 5032–5043, 2018.

- [5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
- [6] Zachary C Lipton, John Berkowitz, and Charles Elkan. A critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.00019, 2015.
- [7] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In International conference on machine learning, pages 1928–1937, 2016.
- [8] Francisco Javier Ordóñez and Daniel Roggen. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1):115, 2016.
- [9] Tim Salimans, Jonathan Ho, Xi Chen, Szymon Sidor, and Ilya Sutskever. Evolution strategies as a scalable alternative to reinforcement learning. arXiv preprint arXiv:1703.03864, 2017.
- [10] John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In International Conference on Machine Learning, pages 1889–1897, 2015.
- [11] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017.
- [12] Mario Srouji, Jian Zhang, and Ruslan Salakhutdinov. Structured control nets for deep reinforcement learning. arXiv preprint arXiv:1802.08311, 2018.
- [13] Felipe Petroski Such, Vashisht Madhavan, Edoardo Conti, Joel Lehman, Kenneth O Stanley, and Jeff Clune. Deep neuroevolution: genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. arXiv preprint arXiv:1712.06567, 2017.
- [14] Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on, pages 5026–5033. IEEE, 2012.
- [15] Duc Trong Tran, Ig Mo Koo, Yoon Haeng Lee, Hyungpil Moon, Sangdeok Park, Ja Choon Koo, and Hyouk Ryeol Choi. Central pattern generator based reflexive control of quadruped walking robots using a recurrent neural network. *Robotics and Autonomous Systems*, 62(10):1497–1516, 2014.
- [16] Duc Trong Tran, Ig Moo Koo, Gia Loc Vo, Segon Roh, Sangdeok Park, Hyungpil Moon, and Hyouk Ryeol Choi. A new method in modeling central pattern generators to control quadruped walking robots. In Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on, pages 129–134. IEEE, 2009.
- [17] Yuan Wang, Kirubakaran Velswamy, and Biao Huang. A novel approach to feedback control with deep reinforcement learning. *IFAC-PapersOnLine*, 51(18):31–36, 2018.

Reengineering the catabolite activator protein system creates a glucose-inducible promoter in *E. Coli*

YASH PERSHAD, RISHABH KAPOOR, ALEX DUNCAN BRADFIELD, TEAGHAN COWLES
STANFORD UNIVERSITY

Type 1 diabetes mellitus (T1D) is an endocrine disorder that affects over 1.25 million Americans. Safe, effective T1D treatments should mimic natural homeostatic functions by continually monitoring blood glucose levels (BGLs) and secreting an amount of insulin appropriate to current BGL. We envision a glucose-sensitive insulin-administration device to replace the function of damaged beta cells in T1D patients. Here, we prototyped a DNA-based glucose biosensor to eventually connect with an insulin actuator. We re-engineered the *E. coli* natural regulatory sequence in the lac operon, which normally increases downstream gene expression under low glucose concentrations, to increase gene expression under high glucose concentrations. Then, we assayed the function of this sensor through GFP expression. Fluorescent assays demonstrated that our construct exhibits glucose inducibility and responds actively to changing glucose concentrations from 0.007 to 0.28 mM. While our results do not suggest that the biosensor operates under physiological conditions, we have created a proof-of-concept DNA-based glucose biosensor. This prototype system has potential to be optimized for physiological conditions, transferred to mammalian cells capable of producing insulin, and ultimately be used for therapeutic applications. Our research represents a new synthetic biology approach to circumvent the current limitations of T1D treatment.

Introduction

Type 1 diabetes mellitus (T1D) is an endocrine disorder that affects upwards of 1.25 million Americans and causes \$14 billion dollars of healthcare expenditures annually. [1] T1D results from an autoimmune attack on beta pancreatic cells, which normally secrete insulin in response to elevated blood glucose levels (BGLs). Without insulin, a peptide hormone that induces systemic cellular uptake of glucose, sustained hyperglycemia and a concomitant lack of cellular energy can result in multi-system failure. [2] Traditional treatment regimens for T1D (Table 1) rely on insulin-replacement therapies to normalize BGL. Since intestinal proteases readily hydrolyze insulin, oral administration is not possible, presenting serious complications for delivery. Furthermore, standard insulin therapy regimens are patient-directed and focus on the administration of insulin in *anticipation* of elevated BGLs (i.e., before a meal).

However, BGLs are subject to variability from diet, activity level, and psychological stress, so patient-directed insulin dosage often does not correspond to actual insulin need. This mismatch often leads to dangerous hypoglycemic episodes from excess insulin administration. Finally, current treatment regimens for T1D are prohibitively expensive. [2][3]

A safer, more effective T1D treatment would mimic the body's natural homeostatic functions by continually monitoring BGLs and secreting insulin only when a patient has high BGLs. This feedback loop would be responsive in real-time and thereby avoid complications from preemptive insulin injections and dosing error. Different therapeutic approaches are at various stages of development, but no single method has been widely adopted as a safe, effective, and cheap alternative to the current standard (Table 1).

In this research project, we aim to leverage genetic engineering to design a glucose-sensitive insulin-administration

device to replace the function of damaged beta cells in T1D patients. Although beta cell response to BGLs is normally mediated by induced exocytosis of insulin, such a complex system could not be readily replicated through genetic engineering. [7] Instead, we envision a DNA-based glucose biosensor that actuates production of insulin (or an insulin-like alternative). Such a device integrated into a patient's somatic cells via viral-vector-based and/or CRISPR/Cas9-based gene therapy would allow modified cells to continually sense BGLs and release appropriate amounts of insulin, thus restoring normal homeostatic functions. CRISPR-mediated therapeutic "knock-ins," where a functional genetic construct is introduced into an organism's genome to compensate for genetic errors in metabolism, have shown promise in preclinical studies for conditions like arginase deficiency.⁸ By reprogramming the patient's own cells, this treatment avoids the complications of infection and immune rejection associated with artificial pancreases and pancreas transplants.

As the first step to creating this synthetic glucose sensor and insulin actuator system, we aim to develop a DNA-based glucose biosensor. Due to the technical difficulties of mammalian cell culture and gene editing, we elected to construct a prototype glucose sensor system in *E. coli* as a proof-of-concept. If the sensor proves effective, it can be transferred to mammalian cells and subsequently optimized for therapeutic application. We selected *E. coli* as our model system for this sensor for the following reasons: (1) *E. coli*'s short doubling time render it a facile, low-cost expression system; (2) stable integration of DNA constructs is readily accomplished through transformation; (3) gene expression in *E. coli* is naturally responsive to changing glucose availability, providing natural regulatory mechanisms that we can repurpose for our sensor.

Although glucose concentrations naturally affect gene expression in *E. coli*, glucose normally serves as a corepressor

rather than inducer of gene expression. Thus, in order to create a glucose biosensor, we aim to re-engineer *E. coli*'s natural systems for increasing gene expression under low glucose concentrations to increase gene expression under high glucose concentrations. Glucose-inducible promoters have previously been engineered in both *E. coli* [11] and *S. cerevisiae* [12], but not with the intention of creating a system that could continually respond to dynamic glucose levels, as would be required for our envisioned *in-vivo* system.

Methods

Plasmid Construction and Validation

Device Design

As *E. coli*'s preferred energy source, glucose represses expression of genes of metabolism for other sugars. [9] For example, the well-characterized lactose (*lac*) operon is subject to positive control by the catabolite activator protein (CAP), also known as the cAMP-receptor protein (CRP). CAP binds to a site upstream of the promoter to recruit RNA polymerase when glucose concentrations are low; thus, transcription of the *lac* operon is repressed when glucose concentrations are high. [10][11]

In order to create an effective glucose biosensor, we aimed to re-engineer the native *E. coli* CAP regulatory system derived from the *lac* operon. We hypothesize that refactoring *E. coli*'s CAP regulatory system by placing the CAP binding site downstream, as opposed to upstream, of the promoter will create a glucose-inducible sensor. We anticipate that increased CAP binding in low glucose concentrations will repress transcription of downstream genes by sterically blocking RNA polymerase progression. Conversely, decreased CAP binding in elevated glucose concentrations will result in increased downstream transcription. Thus, swapping the relative positions of the CAP binding site and promoter should effectively transform CAP from a transcriptional activator to a repressor, which has been demonstrated by previous investigations. [13]

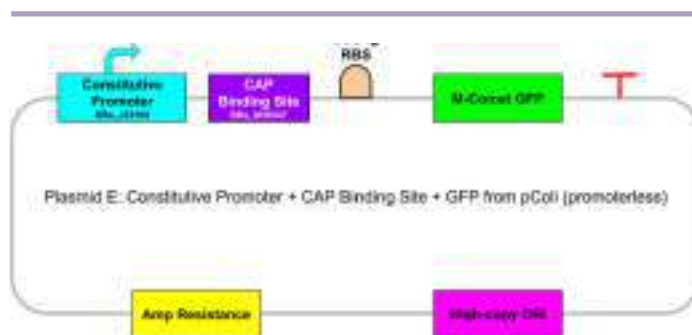


Figure 1. Plasmid E, which represents the novel method for glucose-induced expression. It contains a CAP binding site downstream from a constitutive promoter, which controls GFP expression. We inserted this into the promoterless pColi template plasmid.

1. Experimental Plasmid (Plasmid E)

For our experimental plasmid (Plasmid E), along with placing the CAP site downstream of the promoter, we replaced the natural *lac* operon promoter with a constitutive promoter (BBa_S05450, iGEM) found upstream of many *E. coli* housekeeping genes. Since the natural *lac* operon promoter has a low affinity for RNA polymerase without upstream CAP binding, this promoter would be poorly suited for Plasmid E, which lacks an upstream CAP binding site. Since the constitutive promoter transcribed by default, this configuration will allow us to directly see the effect of the downstream CAP binding site.

We placed the constitutive promoter immediately upstream from a CAP binding site (BBa_M36547, iGEM) [13] in Plasmid E (Figure 1). We elected not to insert a spacer region between the constitutive promoter and the CAP binding site because previous research indicated that the distance between these sequences and the strength of CAP-binding-induced steric hindrance are inversely related; presumably, maximizing

Treatment	Mechanism	Disadvantage(s)
Subcutaneous insulin injection ²	Patient-administered insulin of various durations of action	Pain Lack of patient compliance Disruption of daily routine/quality of life
Insulin pump ²	Electric pump connected to an in-dwelling catheter for continuous administration of slow-acting insulin and fast-acting bolus prior to a meal.	Infection risk Hypoglycemic episodes
Pancreas transplant ⁴	Replace dysfunctional beta cells with pancreatic transplants from a healthy	Immune rejection Operative complications
Artificial pancreas ⁵	Insulin-pump controlled by a continuous glucose monitor	High cost Some models have low portability–utility restrict- More portable models are not fully automated; require finger-prick calibration and food-intake

Table 1. Traditional and developing treatment methodologies for T1D. Established regimens are unshaded, while currently developing methods are shaded in gray.

CAP's repressive activity will lead to greater glucose sensitivity.

We placed the constitutive-promoter-CAP-binding-site complex upstream from GFP in order to assay expression levels through fluorescent intensity. We used the reporter protein GFP as a proxy for insulin because fluorescence enables straightforward tracking of expression levels.

2. Constitutive Promoter Plasmid (Plasmid C)

In order to ensure that this constitutive promoter works and does not itself exhibit a confounding glucose-concentration dependence, we designed a simple positive-control plasmid (Plasmid C) consisting of a constitutive promoter upstream of GFP in the promoterless pColi backbone (Figure 2).



Figure 2. Plasmid C, with a constitutive upstream from GFP. We inserted the constitutive promoter into the promoterless pColi template plasmid.

3. Natural Lac Operon Plasmid (Plasmid N)

Our second positive control plasmid (plasmid N) represents the natural system derived from the lac operon, with glucose-repressed expression. In the event that Plasmid E does not exhibit glucose sensitivity, this control will allow us to eliminate the possibility that the cell has insufficient CAP protein present for regulation of both genomic and plasmid DNA. If the cell has sufficient CAP, this construct should demonstrate an inverse relationship between glucose concentration and fluorescence. This plasmid consists of a CAP binding site (BBa_M36547, iGEM)¹³ upstream of an RNA polymerase binding site (pSB2K4, iGEM) and GFP in the promoterless pColi backbone (Figure 3).

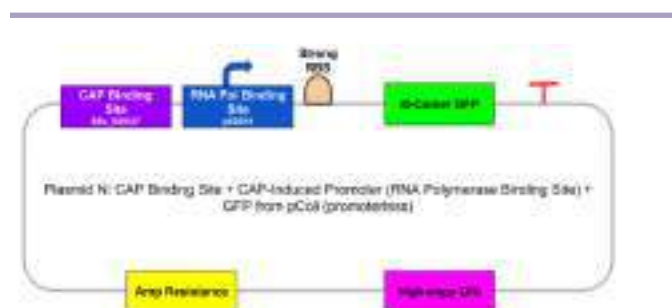


Figure 3. Plasmid N, which represents the natural CAP system upstream from GFP. We inserted this into the promoterless pColi template plasmid.

Device Validation

We obtained the plasmids via Gibson cloning, in which our novel DNA constructs were integrated into a plasmid backbone

using an exonuclease diges't followed by annealing, polymerization, and ligation.¹⁴ Then, we transformed the plasmids into *E. coli* via heat-shock transformation. We selected for transformed cells on LB plates with ampicillin. We picked and resuspended a colony for each plasmid and created a glycerol stock for future experiments. We miniprep'd the transformed *E. coli* to extract plasmid DNA. Then, we executed diagnostic PCR with standard primers placed at the 5' regulatory sequence and in the middle of the GFP coding region for each plasmid. Subsequently, to confirm the plasmid constructs were correctly constructed, we performed gel electrophoresis and sequenced the amplified regions.

Fluorescent Plate Reading

From the glycerol stock, we resuspended bacteria with each plasmid in 10 mL of LB broth or EZ-Rich Medium overnight before each experiment. Before each dynamic range experiment, we transferred bacteria from the culture media to solutions of varying glucose concentrations in a 96-well plate. Our experimental glucose concentrations ranged from 0.002-17.78 mM, with particular concentration ranges determined by the goals of each experiment. We selected this concentration range according to the lac operon's natural glucose dynamic range from 0 to 5 mM and the relevant physiological BGLs from 7 to 15 mM.^{11,12} We chose the upper glucose concentration limit of 17.78 mM in order to simulate the post-meal hyperglycemic conditions in T1D patients.

To prepare fluorescent plates for analysis, we added 800 μ L of EZ-Rich Medium, 16 μ L of glucose solution of various concentrations, and 100 μ L of suspended bacteria to one-mL wells in a 96-well plate. Then, we transferred 200 μ L of the solution from each well to a transparent 96-well plate for measurement of fluorescence and absorbance.

To assay GFP expression, we measured fluorescence via 400-nm excitation and 515-nm emission. To quantify the number of bacteria present, we measured absorbance at 600 nm. From these values, we calculated an OD-600-normed fluorescent value to represent the fluorescence per cell in each condition. For each experiment, we measured these values every 15 minutes for the first three hours, and for experiments measuring response time of bacteria to changing conditions (i.e., glucose concentrations), we continued sampling each hour for nine hours after the initial three. The data were processed, analyzed, and plotted in Matlab.

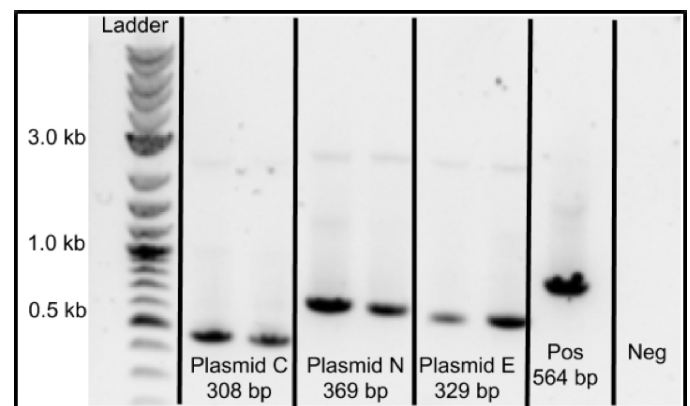


Figure 4. Image of a gel after electrophoresis that shows correct relative lengths of the three plasmids compared to the positive control

Results

Device Validation

The gel electrophoresis confirmed correct length of the region between the 5' regulatory sequence and the middle of the GFP coding region for all three plasmids (Figure 4).

Sequencing analysis corroborated that the constructs were accurately cloned. Therefore, we concluded that the cloning process to construct our plasmid was successful.

Experiment 1: Initial Dynamic Range Fluorescence Test for All Plasmids

We resuspended bacteria with each plasmid in 10 mL of LB broth overnight. For each plasmid, we measured fluorescence over glucose concentrations ranging from 17.78 mM to 0.03 mM using two-fold dilutions and 0 mM to characterize a dynamic range.

Plasmid E potentially demonstrates a dynamic range from 0 mM to 0.28 mM (Figure 5), albeit with overlapping error bars. This dynamic range does not extend to higher glucose concentrations, as the positive association between fluorescence and glucose concentration does not hold above a concentration of 0.28 mM.

Plasmid C exhibits a relatively constant fluorescence level across glucose concentrations, which is consistent with expected constitutive expression. This result suggests that Plasmid E's concentration dependence arises from the presence of the CAP binding site downstream of its promoter rather than from the constitutive promoter.

Plasmid N exhibits a significant decline in fluorescence

between 0.035 mM and 0.07 mM, which is consistent with the lac operon's natural glucose-induced repression. However, this trend does not hold for subsequent increases in glucose concentration. Thus, it is unclear whether Plasmid N exhibits a dynamic range.

Experiment 2: Focused Dynamic Range Fluorescence Test for All Plasmids

Due to the clear lack of a trend in fluorescence beyond 0.56 mM for all plasmids and the high variability in results, we chose to repeat this experiment with a narrower range of seven glucose concentrations (1.11 mM to 0.035 mM with two-fold dilutions and 0 mM).

Within this smaller range of glucose values, the glucose dose-response curve demonstrates clear glucose-induced fluorescence for Plasmid E. For Plasmid C, we see no clear glucose-dependence. For Plasmid N, there is an indication of glucose-repressed fluorescence for lower glucose concentrations, but this negative correlation does not hold for higher values in this range. As in Experiment 1, the relative magnitudes of the normed fluorescence are similar for Plasmids E and C, both of which are higher than Plasmid N; this result is expected, since Plasmid N uses a promoter with less affinity for RNA polymerase.

The results of Experiments 1 and 2 corroborated previous findings and demonstrated evidence of a dynamic range from 0 mM to 1 mM of glucose for the experimental plasmid.⁷ Since Plasmids N and C were meant to troubleshoot Plasmid E and results have confirmed Plasmid E's functionality, we chose to focus future experiments on Plasmid E.

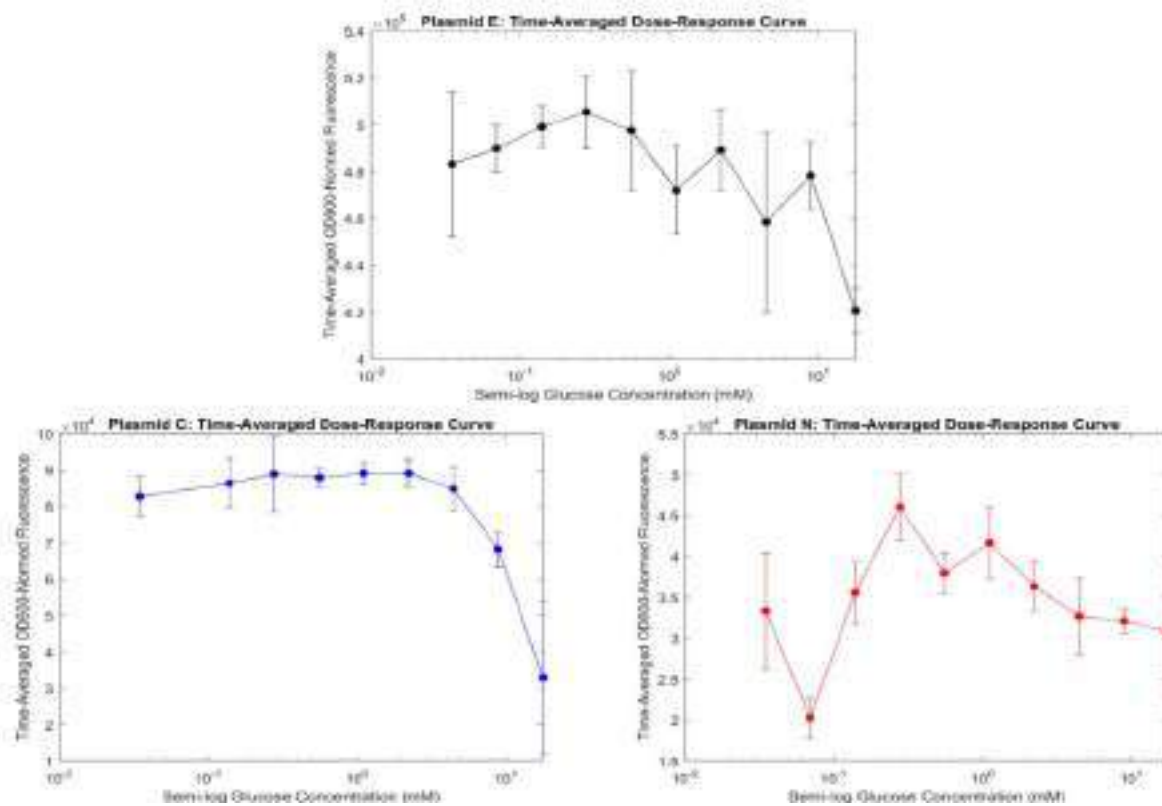


Figure 5. Time-averaged dose-response curves for Plasmids E (black), C (blue), and N (red) from initial time point to three hours with sampling every 15 minutes. Glucose concentrations ranged from 0.035 mM to 17.78 mM with each concentration doubling the previous. Normed fluorescence levels at a glucose concentration of 0 mM (not shown on semi-log plot) fell below the error bars of the fluorescence level for 0.035 mM for Plasmid E. Mean \pm standard error shown for each point.

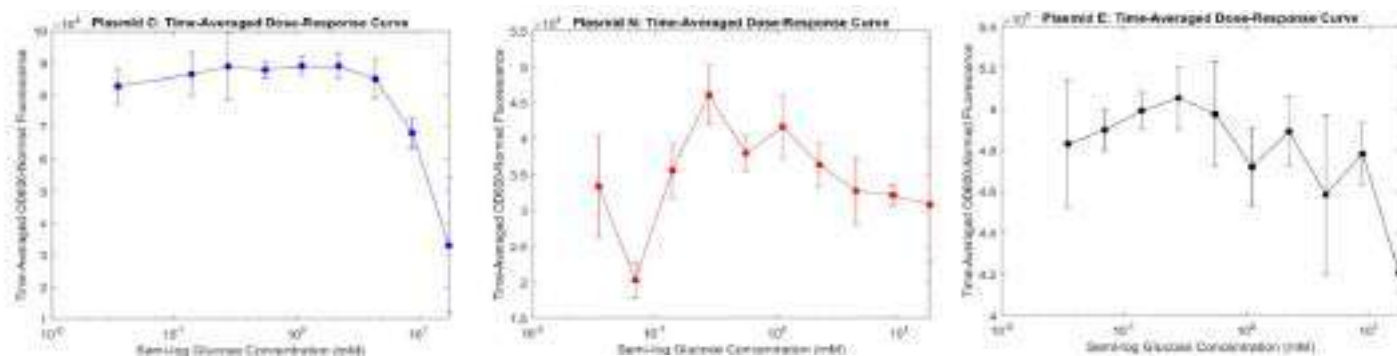


Figure 6. Time-averaged dose-response curves for Plasmids E (black), C (blue), and N (red) from initial time point to three hours with sampling every 15 minutes. Glucose concentrations ranged from 0.035 mM to 1.11 mM with each concentration doubling the previous. Normed fluorescence levels at a glucose concentration of 0 mM (not shown on semi-log plot) fell below the error bars of the fluorescence levels for 0.035 mM for Plasmid E. Mean \pm standard error shown for each point.

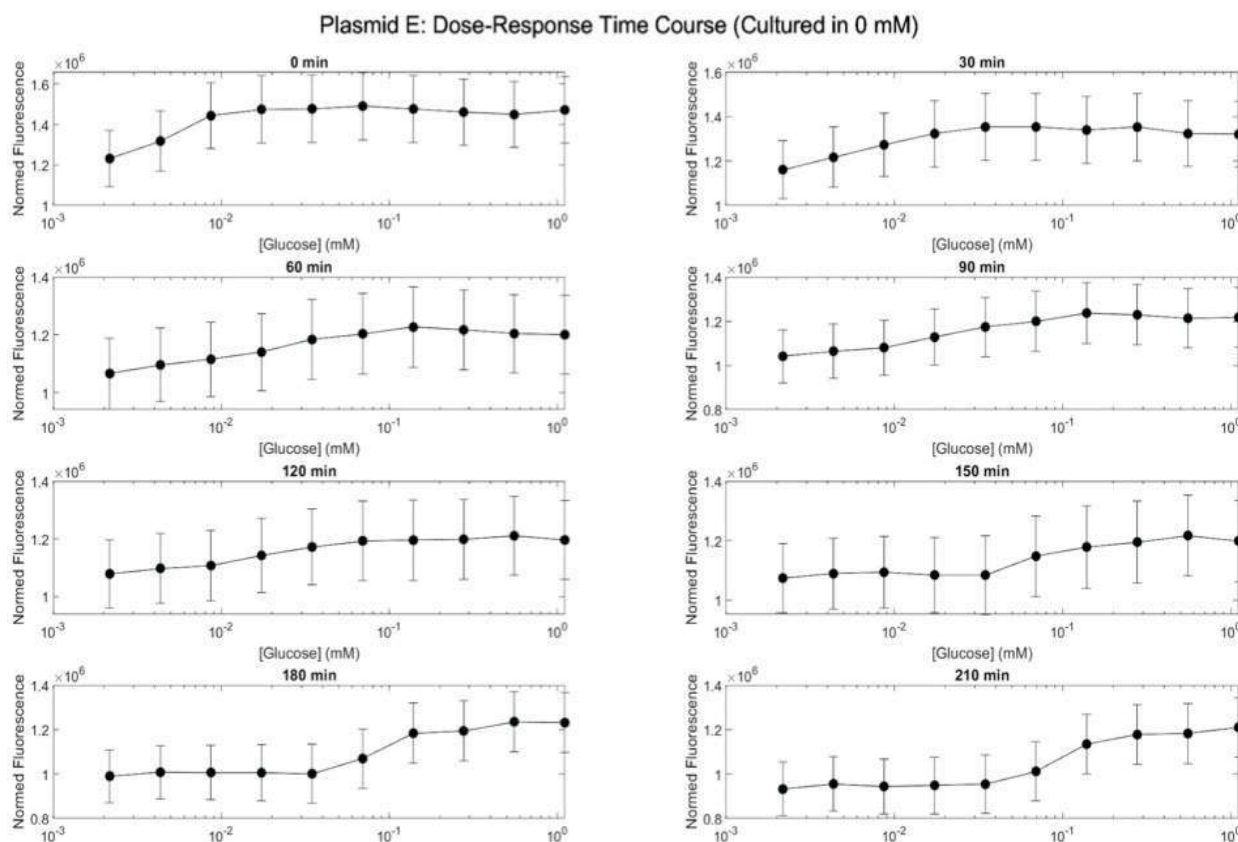


Figure 7. Panel showing glucose dose-response plots for Plasmid E over time from 0 to 210 minutes, with sampling every 30 minutes. Glucose concentrations ranged from 0.002 mM to 1.11 mM with each concentration doubling the previous. Bacteria were cultured in EZ-Rich Medium without glucose. Mean \pm standard error shown for each point.

Experiment 3: Dynamic Range Time Course for Plasmid E Cultured in Glucose-Free Media

Experiments 1 and 2 do not reveal the time course of the sensor's response to changing glucose concentrations, as the LB culture media does not have a defined glucose concentration. Therefore, we performed another dynamic range experiment by culturing the bacteria overnight in EZ-Rich Medium without glucose and transferring them to higher glucose concentrations. This allows us to determine the time course of the sensor's equilibration to an increase in glucose concentration. We tested a greater number of glucose concentrations by performing two-fold

dilutions from 1.11 mM to 0.002 mM, along with 0 mM. By increasing the number of dilutions, we aimed to establish the lower limit of the sensor's dynamic range.

Although the overlapping error bars complicate analysis, the dynamic range began between 0.002 and 0.0035 mM and shifted to 0.007 to 0.28 by 150 minutes (Figure 7). This final dynamic range appears to remain relatively constant after 150 minutes.

Experiment 4: Focused Dynamic Range Fluorescence Test for Plasmid E Cultured in Glucose-Rich Media

We performed a complementary experiment to our last by incubating bacteria transformed with Plasmid E in a glucose-rich

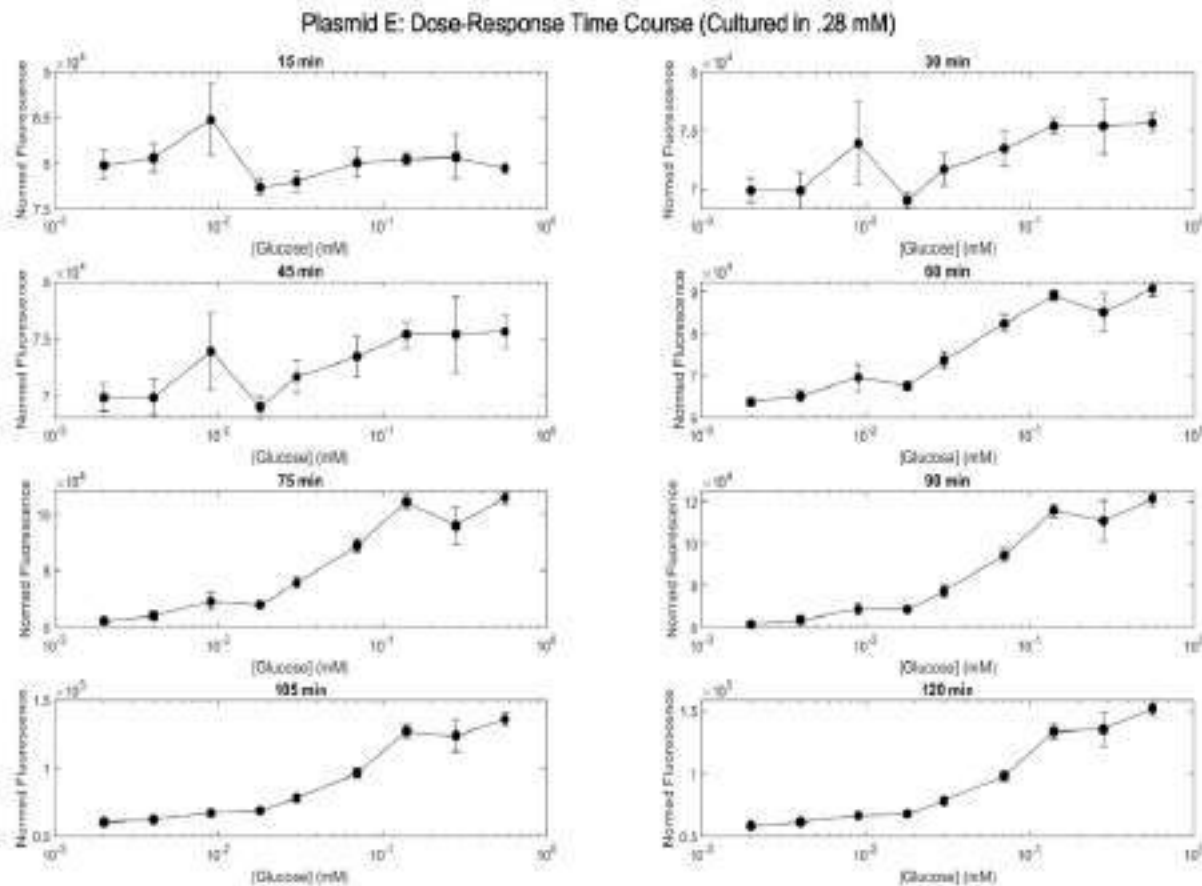


Figure 8. Panel showing glucose dose-response plots for Plasmid E over time from 15 to 120 minutes, with sampling every 15 minutes. Glucose concentrations ranged from 0.002 mM to 1.11 mM with each concentration doubling the previous. Bacteria were cultured in EZ-Rich Medium with 0.28 mM concentration of glucose. Mean \pm standard error shown for each point.

environment (0.28 mM) and transferring them to glucose concentrations ranging from 0.002 mM to 0.56 mM. This allows us to analyze the sensor's equilibration to a decrease in glucose concentration.

After transfer to glucose-rich media, the bacteria began to show glucose-dependent expression at 60 minutes and maintained this glucose-induced fluorescence relationship consistently for all subsequent measurements (Figure 8). This response occurred faster than for the bacteria that incubated in glucose-starved media and also appeared to be more precise, as indicated by the narrower error bars. The sensor appeared to exhibit a clear dynamic range over 0.014 to 0.28 mM, which roughly corresponds to the range determined from Experiment 3.

Discussion

For this project, we sought to create a synthetic DNA-based glucose biosensor in *E. coli* to eventually be paired with insulin-actuator system to treat elevated BGLs for T1D. To do this, we re-engineered *E. coli*'s natural glucose-repressed lac operon to create a glucose-inducible promoter. We designed Plasmid E and hypothesized that placing the CAP binding site downstream from a constitutive promoter would create a glucose-inducible system that could be assayed by expressing GFP. Our experiments aimed to establish a dynamic range for Plasmid E that demonstrates clear glucose-induced fluorescence.

In Experiment 1, we observed glucose-dependent expression at sub-millimolar glucose concentrations, but this

relationship was not present at higher, physiologically relevant levels of glucose. This dynamic range corroborates previous research with the CAP binding site.¹¹ From this experiment, we concluded that the current construct is not glucose-inducible at physiological levels. Experiment 2 corroborated the existence of a dynamic range from 0.007 to 0.28 mM.

Experiments 3 and 4 included lower glucose concentrations, which demonstrated a dynamic range from 0.014 to 0.28 mM. Additionally, our sensor can respond to changing glucose concentrations, which would be necessary for physiological applications. Furthermore, the bacteria respond faster and more uniformly to a decrease in glucose concentrations than to an increase. This difference could alternatively be attributed to irregularities of gene expression and metabolism in the absence of glucose; a follow-up experiment might investigate the response when bacteria are cultured in low but non-zero glucose concentrations and then transferred to higher concentrations.

Our findings suggest that our construct works successfully as a glucose-inducible promoter at low glucose concentrations and responds actively to changing glucose concentrations. However, therapeutic application of our sensor will be limited by its unresponsiveness at physiological BGLs and its slow response time.

If this glucose sensor is to be used as part of a gene-therapy for T1D, these limitations in dynamic range and response time must be addressed. Experimenters might attempt to place multiple CAP binding sites downstream from the constitutive

promoter, which would theoretically amplify the glucose-inducible expression. Additionally, future researchers could vary the distance between the CAP binding site and the promoter; in our construct, we placed a single site immediately downstream. Changing these parameters could potentially shift the plasmid's dynamic range to include physiologically relevant values and quicken response times. Alternatively, researchers might investigate glucose-linked regulatory mechanisms in mammals to create a more responsive and physiologically appropriate sensor.^{16,17}

While existing glucose-inducible promoters have been previously engineered in both *E. coli* 10 and *S. cerevisiae* 11, they were not with the intention of creating a system that could continually respond to dynamic glucose levels, as would be required for our envisioned diabetic treatment. While our results do not prove that the plasmid would operate under physiological conditions and *E. coli* cannot perform post-translational modifications to produce active insulin, we have created a proof-of-concept DNA-based glucose biosensor. This prototype system can be optimized for physiological conditions, transferred to mammalian cells capable of producing insulin, and used therapeutically. Our research represents a new synthetic biology approach to address the current limitations of T1D treatment.

Safety and Security

Our device would be incorporated into a larger glucose-sensitive insulin-administration device for use in T1D patients to monitor and regulate BGLs. This device would then be inserted into the genome of a T1D patient's somatic cells in order to functionally replace damaged beta cells. Our biosensor could not be used for human patients, as currently it would not perform under physiological conditions and cannot adjust glucose levels via an insulin actuator. Thus, an accidental release of the bacteria into a non-laboratory setting would not result in negative consequences. Our bacteria do possess ampicillin resistance and would present problems if they were to be introduced into the world; however, as a BSL-1 hazard, the strain of *E. coli* used is not pathogenic.

Our completely engineered device would, in theory, respond to BGLs by secreting an appropriate amount of insulin into the diabetic patients' bloodstream. A maliciously inclined individual could repurpose our designed pathway to secrete a harmful substance. The aforementioned scenario is highly unlikely, as the effort needed to repurpose our design would be much greater than alternative means of harm. If further experimentation was performed to optimize our system for use in mammalian cells, extensive testing would be necessary before prokaryotic DNA was introduced into a human host. Moreover, the human genome does not naturally encode the CAP protein; so, in addition to transfecting patient cells with the glucose-sensitive insulin device, we would also have to insert a CAP actuator. This construct would also require rigorous safety testing in mammalian tissue cultures and animal models to ensure that the introduction of CAP does not have a deleterious effect on cellular metabolism.

The low risk of adverse effects and potential misuse lead us to conclude that it is ethical to continue to develop our device. However, the cost-effectiveness of investing in our design over other potential diabetic treatments is a necessary consideration. How efficient is allocating money to our device when treatments to manage diabetes already exist? Because current therapies for T1D are prohibitively expensive and not physiologically specific, the medical need for better treatments is extremely high. A synthetic biology approach via a DNA-based glucose biosensor and insulin actuator

offers a potential avenue toward creating better T1D treatments.

Acknowledgements

We thankfully acknowledge Stanford University's BIOE44 (Fundamentals for Engineering Biology Laboratory) instructors, Paul Vorster and Lei Stanley Qi, for their guidance and instruction throughout this research. Additionally, we would like to thank Michael Becich and Samuel Bray, along with the rest of the BIOE44 teaching assistants, for their advice and help.

References

- [1] Zimmet, P., Alberti, K. G., Magliano, D. J. & Bennett, P. H. Diabetes mellitus statistics on prevalence and mortality: Facts and fallacies. *Nat. Rev. Endocrinol.* 12, 616–622 (2016).
- [2] Juvenile Diabetes Research Foundation. Type 1 Diabetes Facts and Statistics. Type 1 Diabetes Facts and Statistics (2009).
- [3] Miller, K. M. et al. Current state of type 1 diabetes treatment in the U.S.: Updated data from the t1d exchange clinic registry. *Diabetes Care* 38, 971–978 (2015).
- [4] NHS Choices. Pancreas transplant. Mayo Clinic (2015).
- [5] Namikawa, T., Munekage, M., Yatabe, T., Kitagawa, H. & Hanazaki, K. Current status and issues of the artificial pancreas: abridged English translation of a special issue in Japanese. *J. Artif. Organs* 21, 132–137 (2018).
- [6] US Food & Drug Administration. The Artificial Pancreas Device System. US Department of Health and Human Services (2018).
- [7] Keizer, J., Magnus G. ATP-sensitive potassium channel and bursting in the pancreatic beta cell. A theoretical study. *Biophys. J.* 1989;56(2):229–42.
- [8] Lee, P. C. et al. Restoring Ureagenesis in Hepatocytes by CRISPR/Cas9-mediated Genomic Addition to Arginase-deficient Induced Pluripotent Stem Cells. *Mol. Ther. - Nucleic Acids* 5, e394 (2016).
- [9] Lawson, C. L. et al. Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* 14, 10–20 (2004).
- [10] Notley-McRobb, L., Death, A. & Ferenci, T. The relationship between external glucose concentration and cAMP levels inside *Escherichia coli*: Implications for models of phosphotransferase-mediated regulation of adenylate cyclase. *Microbiology* (1997).
- [11] Morita, T., Shigesada, K., Kimizuka, F. & Aiba, H. Regulatory effect of a synthetic CRP recognition sequence placed downstream of a promoter. *Nucleic Acids Res.* (1988).
- [12] Cuero, R., Navia, J., Agudelo, D. & Medina, P. Construct of DNA glucose sensor yeast plasmid for early detection of diabetes. *Integr. Obes. Diabetes* 3, 1–9 (2017).
- [13] Czarniecki, D., Noel, R. J. & Reznikoff, W. S. The -45 region of the *Escherichia coli* lac promoter: CAP-dependent and CAP-independent transcription. *J. Bacteriol.* (1997).
- [14] Baeshen, N. A. et al. Cell factories for insulin production. *Microb. Cell Fact.* 13, 1–9 (2014).
- [15] Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA 3rd, Smith HO (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases". *Nature Methods.* 6 (5): 343–345.
- [16] Xie M, Ye H, Wang H, Charpin-El Hamri G, Lormeau C, Saxena P, et al. B-Cell-Mimetic Designer Cells Provide Closed-Loop Glycemic Control. *Science.* 2016;354(6317):1296–301.
- [17] Duan, F. F., Liu, J. H. & March, J. C. Engineered commensal bacteria reprogram intestinal cells into glucose-responsive insulin-secreting cells for the treatment of diabetes. *Diabetes* 64, 1794–1803 (2015).

Applications of Algebraic Topology to Video Compression

JASMINE BAYROOTI & GUNNAR CARLSSON
STANFORD UNIVERSITY

In this study, we applied Algebraic Topology techniques to extract information about the shape of data and applied these insights to the research problem of video compression. Specifically, we applied a computational tool known as Persistent Homology to point-cloud data sets and extracted insights on the data from the induced barcodes. We generalized results from a study of the local behavior of spaces of natural images by Carlsson et al [1] to the study of videos. To this end, we considered an ambient space of 81-dimensional points containing arrangements of 3×3 patches of pixels extracted from the frames within a video. We developed a computational model for the high-contrast dense sub-manifolds of this point cloud and found that these sub-manifolds have the topological properties of a connected bouquet of spheres. The reduction of dimension to a bouquet of spheres could have potential applications for video compression.

Introduction

Topology is the study of properties of spaces and homology is a tool that mathematicians use to characterise the shape of a space. Algebraic Topology is a branch of mathematics that requires background knowledge in subjects including Groups, Rings, and Fields as well as Metric Spaces, Topological Spaces, and Analysis. It works by assigning algebraic invariants, such as a group, to topological spaces. There are a number of ways of manipulating point cloud data to transform the data into a representative topological space. Recently, Carlsson et al [1] at Stanford University have developed a computational software called Plex [2] that enables efficient computations on point cloud data, including calculations of Persistent Homology and barcodes.

The algebraic invariant in Persistent Homology, called Betti numbers, can be thought of as identifying the shape of the data.

For instance, if a point cloud is in the shape of a circle, Persistent Homology would identify that as a 1-dimensional hole in \mathbb{R}^2 as shown in Figure 1. In higher dimensions, there may be n -dimensional holes or spheres embedded in the data. In addition to holes, Persistent Homology is well known for a large class of objects and these characterizations may be used to better understand the shape of the data. For instance, objects like the Klein Bottle shown in Figure 2 are amongst a vast database of objects for which the homology is well understood.

In this section, we give a brief introduction to Persistent Homology and barcodes largely following the Math Review from a JavaPlex tutorial [2] and online paper [3]. Readers interested in Algebraic Topology should refer to Armstrong [4] and to Zomorodian and Carlsson [5] for more details on Persistent

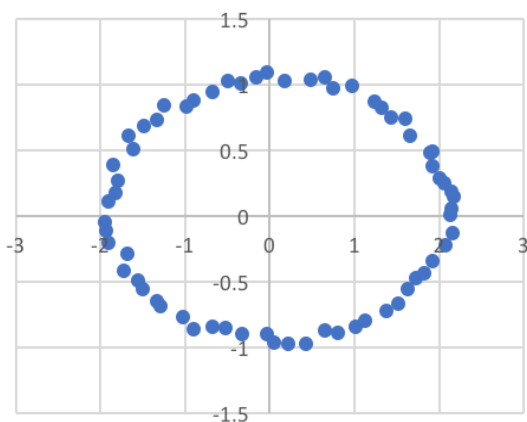


Figure 1. Given a high-dimensional point cloud, it is often difficult to know if linear statistics can be applied. For instance, in this shape the linear trendline would be a misleading representation of the dataset. On the other hand, Persistent Homology would characterize this data with one 1-dimensional hole.

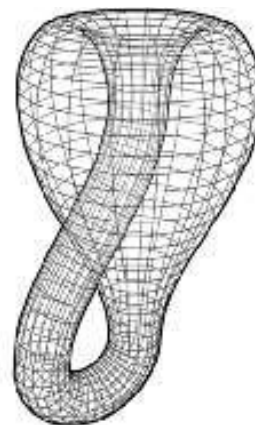


Figure 2. The Klein Bottle is 2-dimensional surface which can be embedded in 4-dimensional space or higher. The shape may be captured in \mathbb{R}^3 as well but then the self-intersection is necessary and adding another dimension creates the Klein Bottle proper without self-intersection.

Homology.

Simplicial Complexes

In a topological analysis, we first replace a set of data points with a family of simplicial complexes to convert the point cloud into a topological space. An abstract simplicial complex is given by the following data.

- A set Z of vertices or 0-simplices
- For each $k \geq 1$, a set of k -simplices $\sigma = [z_0, z_1, \dots, z_k]$, where $z_i \in Z$.
- Each k -simplex has $k+1$ faces obtained by deleting one of the vertices. The following membership property must be satisfied: if σ is in the simplicial complex, then all faces of σ must be in the simplicial complex.

We think of 0-simplices as vertices, 1-simplices as edges, 2-simplices as triangle faces, and 3-simplices as tetrahedrals.

Homology

Betti numbers describe the homology, which one can think of as holes, of a simplicial complex X . The value $Betti_k$, where $k \in \mathbb{N}$, is equal to the rank of the k^{th} homology group of X . $Betti_k$ can be thought of as giving the number of k -dimensional holes and $Betti_0$ is the number of connected components. For instance, the circle has a shape described by $Betti_0 = 1$ and $Betti_1 = 1$. In $\mathbb{Z}/2\mathbb{Z}$, the Klein Bottle has a shape described by $Betti_0 = 1$ and $Betti_1 = 2$, $Betti_2 = 1$.

When one works with a point cloud of data, there is often noise embedded in the dataset. For instance, large amounts of financial data, readings from sensors, and pixels from images, all contain some amount of noise due to a variety of factors. Standard homology of a simplicial complex does not offer a way to deal with that noise. Persistence and barcodes, whose definitions rely on filtered simplicial complexes, are rigorous responses to this problem.

Filtered Simplicial Complexes

A filtration on a simplicial complex X is a collection of sub-complexes $\{X(t) | t \in \mathbb{R}\}$ of X such that $X(t) \subset X(t')$ whenever $t \geq t'$. The filtration

value of a simplex $\sigma \in X$ is the smallest t such that $\sigma \in X(t)$. There are many ways one can generate a filtered simplicial complex, for instance by introducing a metric. We start with a vertex z_0 for $X(0)$ and, define $X(t)$ as all vertices and edges in the larger simplicial complex that are a distance at most t away from z_0 .

Persistent Homology and Barcodes

Given a filtered simplicial complex, those topological features which persist over a significant parameter range of t are considered as signal while short-lived features are noise. Barcodes show holes on an interval, with short intervals corresponding to potential noise and longer intervals corresponding to topological features that persist. Betti intervals help describe how the homology $X(t)$ changes with t . A k -dimensional Betti interval, with endpoints $[t_{start}, t_{end}]$, corresponds roughly to a k -dimensional hole that appears at filtration value $t_{start} < t < t_{end}$ and closes at the value t_{end} . This can be viewed in the form of a barcode graph, as shown in Figure 3.

Using Algebraic Topology to Compress Optical Images

Summary of Optical Image Study

In the paper *On the Local Behavior of Natural Images* [1], Dr. Carlsson et al showed that a large subset of 9-dimensional data from natural images lies on the surface of a Klein Bottle, as observed through the “three circle” model (see p. 8-9 of [1] for details). In the paper *The Ring of Algebraic Functions on Persistent Barcodes* [6], this result was extended with the goal of compressing storage of optical photos. In this study, we verify the findings in [1] on a different set of images and generalize the approach to videos.

Procedure for Image Analysis

Carlsson et al [1] performed their analysis on photos from a database of images constructed by H. van Hateren [7]. In our study, we randomly select different optical photos from a private photo library to determine whether or not the results continue to hold. Given the different set of photos, we follow the steps taken in [1]:

Step 1: Randomly select approximately 50,000 size 3×3 patches of pixels from a set of unconnected grayscale images in the database. Each 3×3 patch corresponds to a 9-dimensional vector x where $x = (x_1, x_2, \dots, x_9) = [I_{11}, I_{12}, I_{13}, I_{21}, I_{22}, \dots, I_{33}] \in \mathbb{R}^9$ where I_{ij} is the intensity of the i_j th pixel.

Step 2: Take the natural logarithm of each coordinate to obtain $x = (\log x_1, \log x_2, \dots, \log x_9) \in \mathbb{R}^9$. According to Weber’s Law, there is an inverse relationship between ambient illumination and human sensitivity to light. Therefore, the ratio of $\frac{dx}{x}$ is constant for a wide range of luminances (see [8] for more details).

Step 3: Compute the D-norm of each vector to obtain a measure of contrast of a patch. The D-norm is calculated by summing the differences between the vertically and horizontally adjacent neighbors in a 3×3 patch and then taking the square root. The D-norm is defined in this way because it is commonly believed that regions of photos with high contrast convey the most significant content of a scene (p. 3 [8]). If the 9-dimensional patch is x , the D-norm is given by $\sqrt{x^T D x}$ where D is a positive definite symmetric matrix shown below (see [8] for more details on derivation). For instance, to obtain the 5th row of the matrix in Figure 4, we take $(x_5 - x_2) + (x_5 - x_4) + (x_5 - x_6) + (x_5 - x_8) = 4x_5 - x_2 - x_4 - x_6 - x_8$.

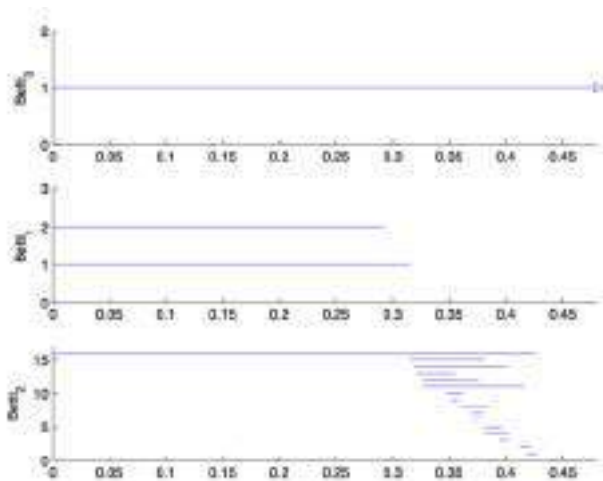


Figure 3. This is an example of a barcode sequence for a point cloud corresponding to a Klein Bottle from figure 2, which confirms that the Betti numbers are $Betti_0 = 1$ and $Betti_1 = 2$, $Betti_2 = 1$. The short barcodes are noise.

$$D = \begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 3 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 3 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 2 \end{bmatrix}$$

Step 4: Select the patches with a D-norm in the top T percent of the entire sample. This is done to maintain the high contrast patches since they follow a different distribution than low-contrast patches and contain the most important information about an image (p. 3 [8]). The resulting point cloud is denoted X_T .

Step 5: Subtract from each vector the average of its coordinates, to get $x = (x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_9 - \bar{x})$ where $\bar{x} = \Sigma x_i$. This reduces the dimension from nine to eight because the sum of the new points will always be zero, meaning if we know eight of the values, the ninth one is determined. Next, divide by the D-norm to normalize the selected vectors. This makes each vector have a unit length so that they all lie on a unit sphere in R^8 . Note that the unit sphere is given by the equation $x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2 + x_8^2 = 1$. Therefore, knowing seven of the coordinates determines the eighth. Hence, the point cloud lies on a 7-dimensional sphere in R^8 .

Step 6: Filter out the outlier points in the remaining point cloud using the kth nearest neighbor density function $p_k(x)$. Select p percent of the points whose Euclidean distance to their kth nearest

neighbor are smallest. This gives a dense subset of the point cloud denoted by $X_T(k,p)$. A small choice of k results in a local density estimate while a larger k value provides a more global estimate.

Step 7: Use the resulting dense point cloud $X_T(k,p)$ as input for topological analysis. This involves running JavaPLEX to create the filtered simplicial complexes and find the corresponding persistent homology and barcodes described in Section 1 for the point cloud. Examining the barcodes will show patterns or submanifolds in the dense point cloud $X_T(k,p)$.

Results for Image Analysis

We found that the barcode representation of persistent homology in Figure 8 as well as the 2-dimensional cross sections (Figures 9-11) corroborate the results from *On the Local Behavior of Natural Images* ([1], p. 5-8) when run on a random assortment of optical images including Figures 5-7 and more.

Discussion for Image Analysis

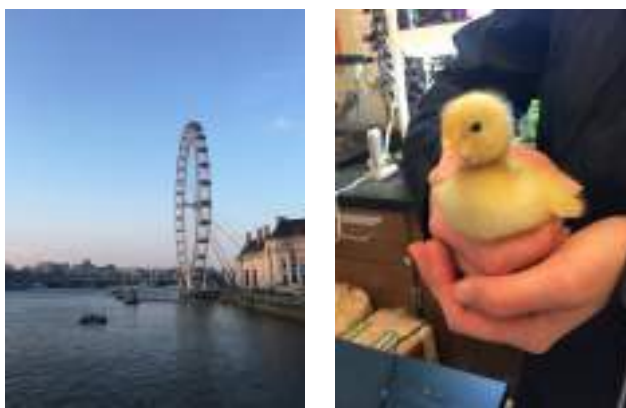
These results and their corresponding homology aligns well with the study by Carlsson et al [1]. In particular, we also observed the three circles in the (e1, e2) plane, (e1, e3) plane, and (e1, e5) plane of the Three Circle Model of the Klein Bottle. This indicates that their conclusion that the high contrasting and dense sets of the original patches lie largely on the surface of a Klein Bottle may apply to other datasets beyond the Van Hateren database. This conclusion makes the compression of images more efficient since the dense subset lies on a 4-dimensional shape rather than a point cloud in 9 dimensions [6], potentially allowing up to a 50% reduction of storage on dense subsets.

Using Algebraic Topology to Compress Optial Videos Procedure for Video Analysis

For the video project, a similar method was followed with a few key differences. The main difference is that a video is a collection of frames or pictures that are taken in linear time and are consequently similar to neighboring frames. To address this difference, we assign consecutive frames to groups and look for areas of large variance or contrast in each group. The steps for the method are detailed below:

Step 1: Process the video in grayscale and divide all of the scenes in the video into groups of nine consecutive frames. For videos with a frame rate of 30 frames per second, as used in this study, nine frames corresponds to approximately 330 milliseconds of video footage.

Step 2: Construct a point cloud by randomly picking 900 3 x



Figures 5, 6. Sample optical images from the database.

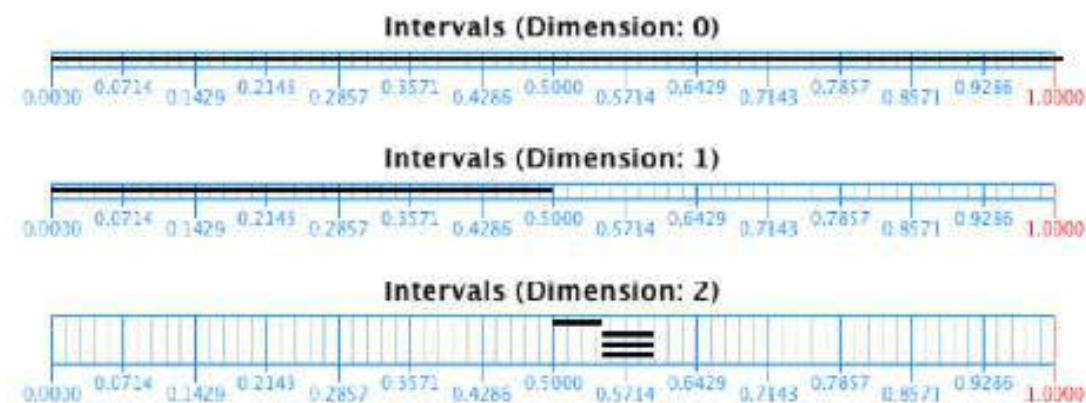


Figure 7. Barcode results from optical image analysis.

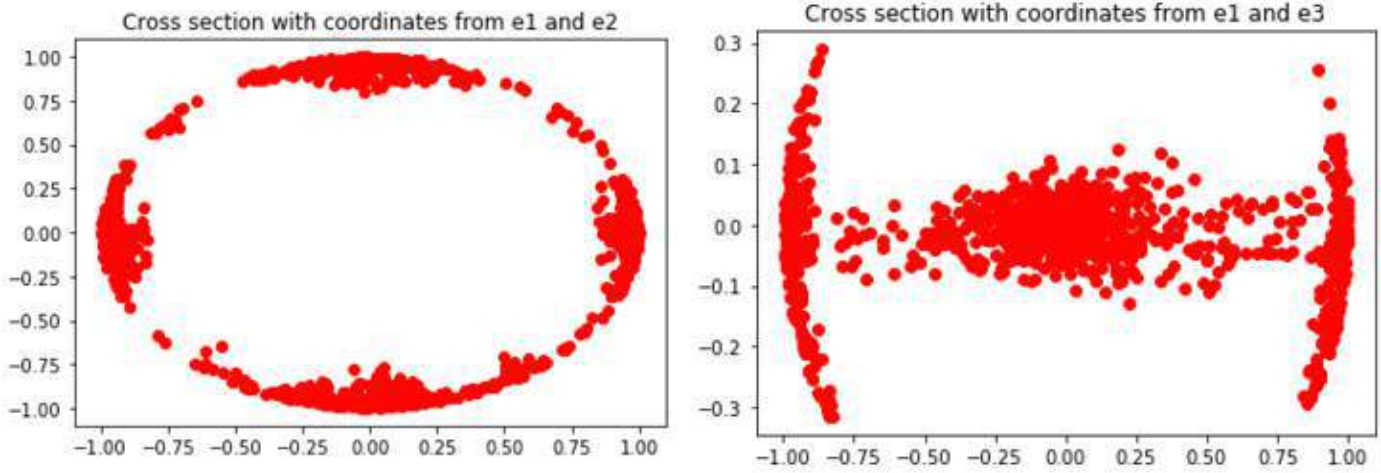


Figure 8. Representations of the shape of data from two different cross sections for optical images with $k=15$, $p=20\%$ and $T=10\%$, $X_{0,1}(15, 0.2)$. These two views are part of the Three Circle Model.

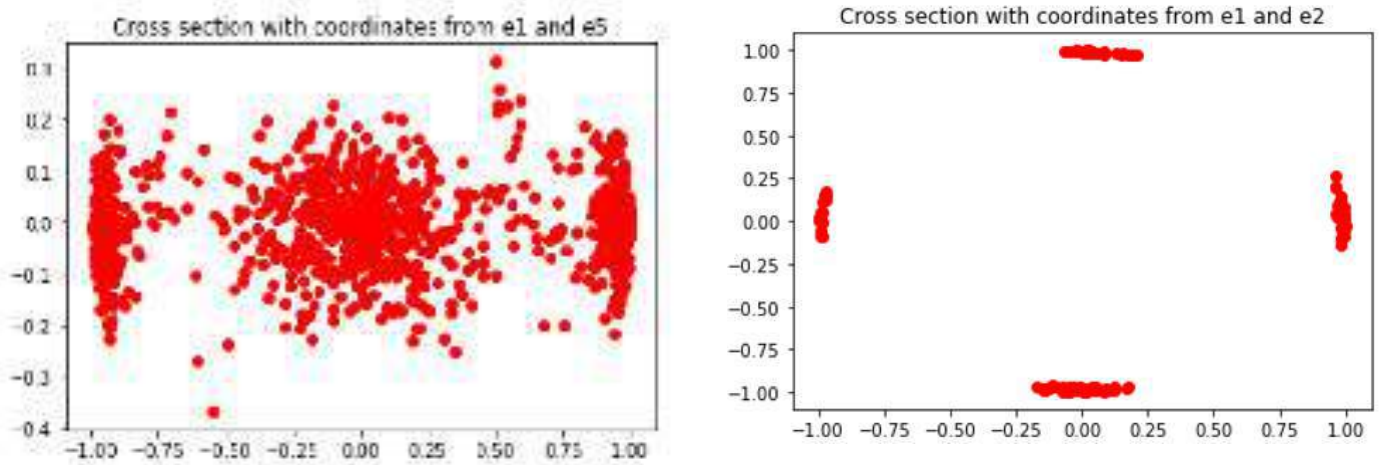


Figure 9. A representation of the shape of data from a third cross section for optical images with $k=15$, $p=20\%$ and $T=10\%$, $X_{0,1}(15, 0.2)$. This is the third part of the Three Circle Model.

Figure 10. A representation of the shape of the data for optical images with $k=300$, $p=3\%$, and $T=10\%$, $X_{0,1}(300, 0.03)$. This shows the four dense corners of the circle.

3 patches of pixels in the same location across all nine frames in the group. As in the picture study, take the natural logarithm of each coordinate of the 3×3 patch to obtain a 9-dimensional vector x , where $x = (x_1, x_2, \dots, x_9) = [I_{11}, I_{12}, I_{13}, I_{21}, I_{22}, \dots, I_{33}] \in R^9$. By considering the 3×3 patches from all nine frames in a group, obtain points in R^{81} .

Step 3: Calculate the D-norm for each point by averaging the D-norms of the nine 9-dimensional vectors from each frame in the group. If the standard deviation of the D-norms from the group is small, then the contrast in that region of the group is small and should be eliminated. Therefore, the point is kept provided that the standard deviation among the nine D-norms is larger than a specified threshold value to ensure that there is significant change within the group itself.

Step 4: In addition to including patches where there is high variation among the nine frames from step 3, we select the patches with a D-norm in the top T percent of the entire sample to produce the point cloud X_r . This ensures that we consider areas of high contrast within the individual frame.

Step 5: Subtract from each vector the average of its coordinates

to reduce the dimension from 81 to 80. Next, divide by the D-norm to normalize the selected vectors. This makes each vector have a unit length so that they all lie on a unit sphere in R^{80} . Note that the unit sphere is given by the equation $x_1^2 + x_2^2 + \dots + x_{80}^2 = 1$. Hence, the point cloud lies on a 79-dimensional sphere in R^{80} .

Step 6: Filter out the outlier points in the remaining point cloud using the k th nearest neighbor density function $p_k(x)$. Select p percent of the points whose Euclidean distance to their k th nearest neighbor are smallest. This gives us a dense subset of the point cloud denoted by $XT(k, p)$. A small choice of k results in a local density estimate while a larger k value provides a more global estimate. Because we have many points in high dimensions, finding the k th nearest neighbor is computationally expensive and is an active area of research. The brute force approach runs in $O(NDk)$ where N is the size of the training set, D is the dimension of each point, ie 79, and k is the number of neighbors to take. We use the KDTree algorithm from the SKLearn package in SciPy. This is a C package that is available within Python and is very efficient as it runs in $O(\log N)$ for searching.

Step 7: Run a topological analysis on the resulting dense point cloud $X_T(k,p)$ to look for patterns or submanifolds through an examination of the barcodes. Visualize the shapes using 2-dimensional cross sections with Python's matplotlib.

Results for Video Analysis

We obtained the following results on a grayscale video. See Figures 11-15.

Discussion for Video Analysis

The persistent barcodes corresponding to dense subsets $X_T(k,p)$ of a point cloud obtained from video footage appears to lie on a surface that is topologically homeomorphic to a bouquet of spheres.

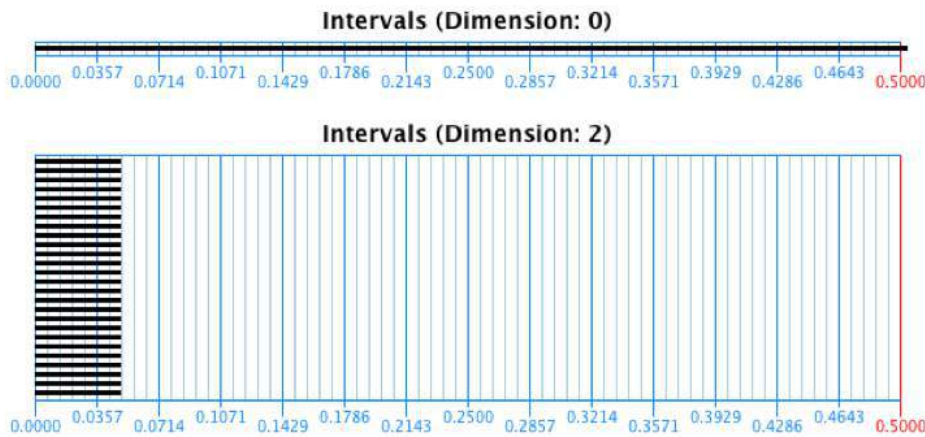


Figure 11. These are the barcode results of the analysis. Since there is only one interval in dimension 0, this means that the surface is connected. Note that there are no holes in dimension 1 and there are a number of distinct holes in dimension 2. 2-dimensional holes are topologically equivalent to spheres. Therefore, the barcodes suggest a shape called a "bouquet of spheres" with the number of intervals n in the second dimension being the number of spheres attached to a common point.

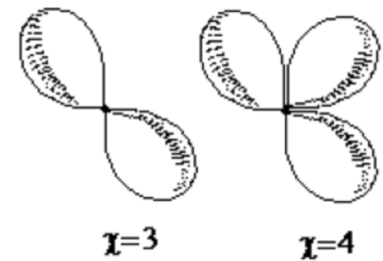


Figure 12. Two representations of bouquets of spheres where $n = 3$ and 4.

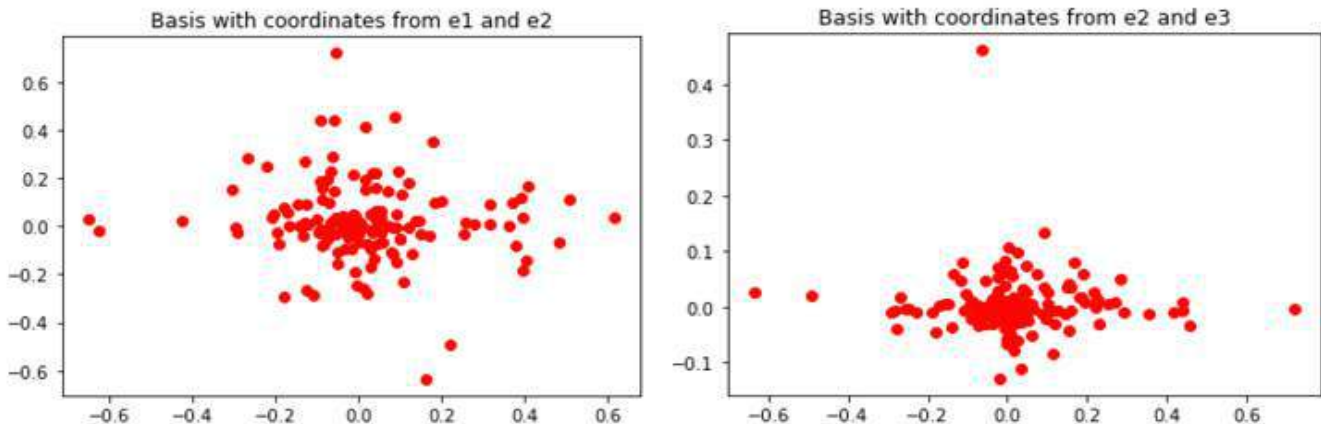


Figure 13. Representations of the data from an optical video with $k=15$, $p=10\%$, and $T=20\%$, $X_{0.2}(15, 0.1)$.

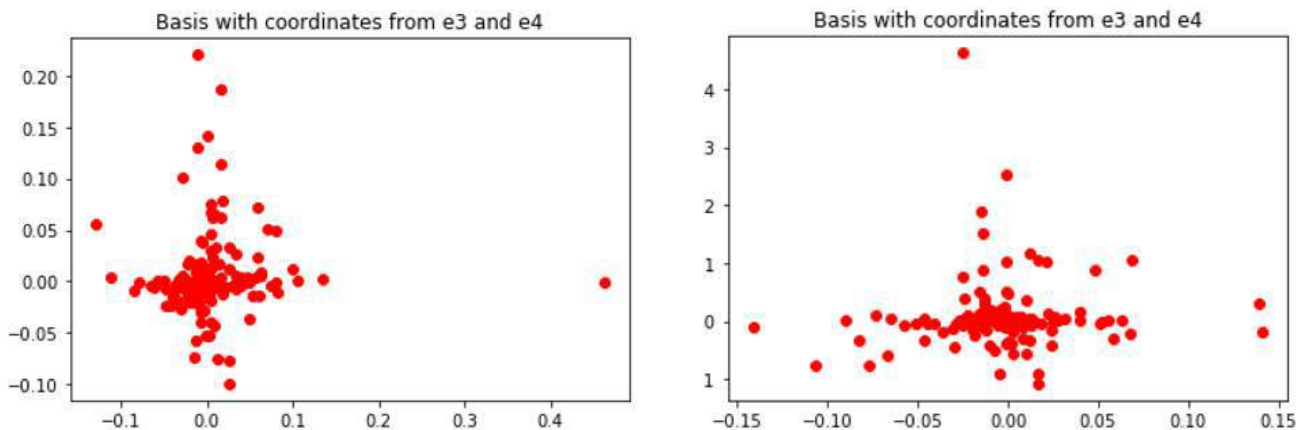


Figure 14. Representation of the data from an optical video with $k=15$, $p=10\%$, and $T=20\%$, $X_{0.2}(15, 0.1)$.

Figure 15. Representation of the data from an optical video with $k=300$, $p=10\%$, and $T=20\%$, $X_{0.2}(300, 0.1)$.

The bouquet of spheres is a surface for which the homology is well understood with holes only in dimension 2. Each of the spheres in the bouquet is a 2-dimensional surface. By preprocessing the video frames, one can potentially transform each pixel in the original 81-dimensional space to a 2-dimensional point on an appropriate sphere. For instance, one could map an 81-dimensional object to a 3-dimensional object, the first component of which identifies which sphere from the bouquet it lands on while the other two components specify the location on the sphere. Since this transformation may be performed on a dense subset of the video, there is significant potential for compression of data. The next steps in this study would be to repeat the analysis on additional video footage to confirm the shape of dense subsets and also to define and test a transformation for compressing the data without losing image quality.

Conclusion

In this study, we applied Algebraic Topology to gather information about the shape of data and used these insights on the research problem of the local behavior of natural videos. We found that the persistent barcodes corresponding to the high-contrast and dense subset of the original 81-dimensional point cloud appeared to lie on a surface that is homeomorphic to a connect bouquet of spheres, which is a topological surface whose homology is well understood. The lower dimensional results could be potentially useful for video compression just as the resulting Klein Bottle shape from the study of local behavior of natural images by Carlsson et al [1] was used for the compression of images. This application of Algebraic Topology shows the great potential in the field for new discoveries and approaches to important research problems.

Acknowledgements

I would like to thank Dr. Carlsson for all of the meetings, emails, skype sessions, and advice he gave me over the course of this research project. It has been a very rewarding process for me personally and it has motivated me in my further research. I would also like to thank my independent math research instructor, Ms. Keshena Richardson, for her guidance, advice, and many valuable comments and Dr. Andrew Ying for his very helpful feedback and editing. Finally, I would like to thank my parents for their support throughout my education.

References

- [1] Carlsson G., Ishkhanov T., De Silva V., Zomorodian A., On the local behavior of spaces of natural images, *International Journal of Computer Vision* 76, 1 (2008), pp.1-12.
- [2] Adams H., Tausz A. *JavaPlex Tutorial*. 2017
- [3] Ghrist R., *Barcodes the Persistent Topology of Data*. *Bull. Amer. Math. Soc.* 45 (2008), 2007, pp. 61-75
- [4] Armstrong, M. A. *Basic topology*. Springer, 2011.
- [5] Zomorodian A., Carlsson G. *Computing Persistent Homology*. *Discrete Computational Geometry*, 33:249-274. 2005
- [6] Adcock A., Carlsson E., Carlsson G., *The Ring of Algebraic Functions on Persistent Barcodes*. 2013.
- [7] van Hateren J. and van der Schaff A., *Independent Component Filters of Natural Images*

- Compared with Simple Cells in Primary Visual Cortex, *Proc. R. Soc. London, B* 265 (1998), 359-366.
- [8] Lee A.B., Pedersen K.S., Mumford D. The non-linear statistics of high-contrast patches in natural images. *International Journal of Computer Vision* 54, 1-3(2003), pp. 83-103

Light and Temperature on Reef Herbivory: Effects of Algal Growth Conditions on Grazing Intensity by Juvenile and Adult Sea Hares

OSANNA DRAKE

STANFORD UNIVERSITY

Herbivore-algae interactions play a dominant role in reef health and stability. As climate change progresses, light exposure and sea surface temperature are predicted to increase. This study examines how herbivores respond to climate change-induced impacts on prey algae using the spotted sea hare (*Aplysia dactylomela*) and the red alga *Laurencia intricata*. Algae was cultured under increased light and temperature treatments then presented to adult and juvenile sea hares. The mass of algae consumed from each treatment by each age group was determined. It was found that algae grown under high temperature was eaten significantly less than algae grown of any other treatment. Adult sea hares ate significantly more per capita than juveniles and treatment preference was not significantly different between the two age groups. These results indicate that increasing sea surface temperature will likely have an impact on reef herbivory and macroalgal abundance and merits further study.

Introduction

The oceans and climate are currently experiencing change at an alarming rate, and their health and stability are projected to deteriorate for many decades to come[1]. It is essential to study the effects of these changes as they concern the welfare and future of the planet as well as its inhabitants[2].

Coral reefs are crucial, as they are locations of immense biodiversity and high ecological value[3]. They support intricate food webs and provide habitat for many organisms, including charismatic flagship species such as sea turtles and whales. Reefs also provide ecosystem services through jobs, food, tourism, recreation, cultural value, and storm protection totaling over AU\$6.6 billion annually[4]. Thus, it is in the planet's and humanity's best interest to prioritize reef preservation during this pivotal time.

It has been well documented that coral reefs are under threat from climate change, as rapid ocean acidification coupled with increasingly frequent bleaching events is steadily eroding these calcifiers[5]. The complexity of physical, chemical, and biological variables involved in reef structure make it difficult to predict how these climatic changes will manifest. One topic that has spurred much controversy is the way climate change is predicted to impact algae[6]. It is well known that one of the dominant factors in reef health and productivity is the coral/macroalgae relationship. Herbivorous reef fishes and invertebrates graze heavily on macroalgae and control its growth in the process. In their absence, or in the presence of excess nutrients, algae will overgrow the coral, blocking incoming light and impeding coral photosynthesis and growth. Additionally, as macroalgae overtakes hard substrates, coral larvae are unable to settle and recruit, initiating a positive feedback loop eventually causing a phase shift from coral reef to macroalgae dominated benthos[7]. Phase shifts are often triggered by coral bleaching and mass mortality events (when the reef is already at a weakened state) and are exceptionally

difficult to recover from. Algal phase shifts not only reduce the reef's aesthetic beauty and tourism value, but also lead to biodiversity loss, fishery collapse, and diminished shoreline protection. Thus, developing an understanding of the processes and precursors to phase shifts has substantial implications in multiple sectors[8]. Herbivory maintains reef health by preventing phase shifts from occurring[9], so it is worthwhile to explore how climate change may impact it. This paper seeks to address this issue by testing how herbivory is altered when presented with macroalgae grown under future climate conditions.

The herbivore selected for this study was *Aplysia dactylomela* (Rang, 1828), the spotted sea hare. Sea hares, a clade of gastropod mollusk, fill an important niche by feeding off algae containing chemical toxins not grazed on by other herbivores[10]. *A. dactylomela* is a cosmopolitan species, found across the world in tropical waters and reefs[11]. This species' broad distribution and key role in controlling toxin-laden algae makes *A. dactylomela* globally relevant and impactful, and thus an ideal choice for this experiment.

A. dactylomela feeds primarily on the red macroalga *Laurencia intricata*, sequestering the secondary metabolites to become unpalatable themselves[12]. *A. dactylomela* also uses *L. intricata*'s pigments to produce a purple ink believed to confuse the sensory organs of would-be predators[13]. The health of the alga determines the amount of chemical defenses it can produce as well as its nutritive value. Thus, as survival of *A. dactylomela* is affected by ability to grow large enough to reproduce while avoiding predation, there is an evolutionary advantage to feeding on algae that is healthier and more nutritious. In fact, several studies [10, 14] have shown that *Aplysia* sea hares are indeed able to detect and selectively feed on algae that is more nutritious. Given this, it follows that *A. dactylomela* will, when all else is constant, preferentially graze on the healthiest, most nutritious algae available.

Future climate models predict the Great Barrier Reef will experience elevated sea surface temperature, as well as increased light exposure during El Niño events: periods of weakened trade winds and altered circulation [15]. These stressors may impact the health and nutritional content of reef macroalgae, which would be apparent in the feeding preferences of herbivores, in this case *A. dactylorella*. By examining how herbivory changes when sea hares are presented with food sources similar to what they will encounter in the future, we can gather information on what reefs may begin to look like as climate change progresses.

It has been shown that light and temperature are two of the most important factors in determining growth rates of algae in the genus *Laurencia* [16] and thus, certain elements of the algae's physiology will likely be altered along with rate of growth and may influence sea hare feeding preference. A study [17] on growth rates of macroalgae under different conditions found that both temperature and light will negatively alter growth rate as well as pigmentation and nutritive value when altered from the natural optimal value. While this does not indicate how this algae will be grazed by herbivores, it would suggest that as a result of elevated light and temperature it is less healthy overall. However, a study [18] in 2000 suggests the opposite, finding that red abalone exhibited higher growth rates when fed algae cultured under higher light conditions, likely indicating an enhanced nutritional content. One study [21] in the Bahamas used a naturally occurring Nitrogen gradient to study the relative impact of N and herbivory (by *A. dactylorella* and others) on macroalgae. They found that *A. dactylorella* grazed more on *L. intricata* grown in high N conditions, presumably due to this elevated nutritional value, and that herbivory was the dominant factor dictating macroalgal species

composition by consuming certain algae to their preference. However, they did not significantly alter total macroalgae biomass and subsequent coral cover, as that was attributed largely to N content.

One study [19] found that dietary preference in limpets and chitons is largely determined by the ability of the animal's mouth parts to accommodate a given algae's structure and texture. For example, the periwinkle snail *Littorina littorea* possesses a large number of small sharp teeth with high surface area and primarily grazes on soft filamentous algae and microalgal films. Conversely, another periwinkle, *Littorina obtusata*, has fewer, stronger, and blunter teeth, and is able to prey on tough macrophytes [19]. A more general study on herbivory [20] looking at isopods and amphipods concluded that dietary preference is a result of either 'attractiveness', here loosely meaning health and nutritional value, or 'edibility'. The paper went on to elaborate that food choice in a given herbivore is dominated by food quality or ease of eating (rarely both), and is influenced by a variety of factors including habitat, physiology, predation pressure, size, and food source availability. To illustrate, the *Idotea baltica* isopod chooses hard, branched algae which, although more difficult to consume, is highly nutritive [20]. Meanwhile *Ampithoe valida*, an amphipod, favors soft, supple bladed algae because of this morphology and the fact that it is a more prevalent food source [20]. Additional elements can also come into play, such as a rigid food alga doubling as wave protection and predator defense for *I. baltica* [20]. While a sea hare is quite different from either an isopod or amphipod, this relationship has been shown to persist: the type of algae a sea hare species predate upon is often linked to the complexity of its radula, with highly specific radulae indicating a specialist,

whereas more modest, uncomplicated radulae typically belong to generalists [22]. In a study regarding the related *Dolabella* sea hares, herbivory was determined largely by the softness of the algae, for in every case in which they were presented with two food choices, they chose the softer [23].

Still other studies [24] have deemed these oversimplifications, and emphasized the notion that food choices cannot always be reduced to simply energy, texture, or protein content. Instead, herbivores moderate their diet to predate on a variety of primary producers to achieve maximum nutritional complementarity [24]. In other words, they prefer a well balanced diet.

The 6cm long *Stylocheilus striatus* sea hare is frequently associated with cyanobacterial blooms, predated on the microalgae. It in turn is targeted by a carnivorous nudibranch (another variety of sea slug), with each nudibranch consuming more than 2 sea hares per day [25]. Predation of sea hares, and the resultant reduction in herbivory, lead to a 50% increase in cyanobacteria biomass [25]. In this case, smaller sea hares were 22 times more likely to be eaten than large individuals [25] so it is worth noting that juveniles may behave differently from adults due to predator avoidance behavior which could lead to differences in food preference. While *Aplysia* sea hares reach sexual

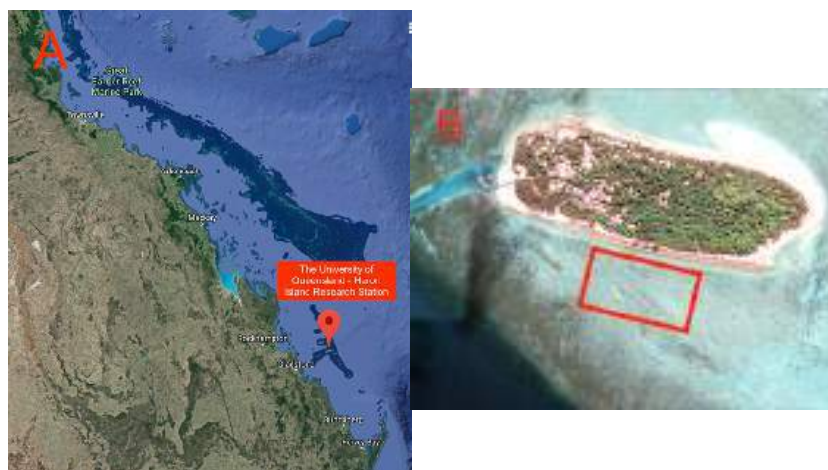


Figure 1. A) Satellite image showing location of Heron Island Research Station (HIRS). HIRS is located on Heron Island (23.442°S, 151.914°E) in the southern Great Barrier Reef. It is operated by the University of Queensland and is the reef's oldest and largest marine station. The Great Barrier Reef Marine Park is shown in dark blue with a red icon designating Heron Island; **B)** Aerial view of Heron Island. Experiments were conducted in HIRS facilities, with organisms and seawater collected on the southern reef flat within the Scientific Research Zone (rectangle outlined in red).
Image source: Google Earth, 2018



Figure 2. *Aplysia dactylomela*, the spotted sea hare, was the herbivore chosen for this study. Arrows show size difference of adult (bottom arrow) and 15 cm juvenile (top arrow) *A. dactylomela*. During a 3d collection period, two adults and four juveniles were found and moved to a large flow-through tank containing rocks from the reef flat and allowed to acclimate for 7d before participating in feeding trials. All were fed a diet of fresh *Laurencia intricata* macroalgae each day.



Figure 3. Experimental setup for algae treatment tanks. Twelve 125L flow-through tanks arranged beneath a shade screen to equalize natural light exposure. Tanks were randomly assigned one of four treatments: Control (C), increased light (L), increased temperature (T), and both increased light and temperature (B). There were three replicate tanks for each treatment. Tanks were filled with *Laurencia intricata* macroalgae and cultured for 5d under treatment conditions. Water for tanks was pumped directly off the reef flat at Heron Island Research Station where the algae was collected. Tanks and equipment were monitored daily to ensure proper functioning.

maturity within several months and spend half their life as adults, the juvenile phase is also significant [26] and thus should be included in analysis. There is evidence that some *Aplysia* species change their preferred algae as they mature [27], but understanding the evolution of dietary preference across life stages is largely unknown [28]. In this experiment, *Laurencia intricata* was exposed to increased temperature and light and the impact on herbivory by the sea hare *Aplysia dactylomela* was measured. Additionally, the impact on herbivory of adult and juvenile sea hares was compared. This study aims to address the following three questions:

1. Will *Aplysia dactylomela* consume *Laurencia intricata* grown in high temperature and/or light differently than that grown in control (present day) conditions?
2. Will adult and juvenile *Aplysia dactylomela* consume different amounts of *Laurencia intricata*?
3. Will adults and juveniles differ in preference between *Laurencia intricata* treatments?

As shifts in herbivory can have widespread impacts [29], answering these questions may shed light on the future of reef community structure, and thus improve management techniques and preparation for climate change.

Methods

Overview

This experiment was conducted at Heron Island Research Station on the Great Barrier Reef (Fig. 1A), and sea hares and algae were collected on the southern reef flat (Fig. 1B). Tides are semidiurnal, and a reef crest separates the flat from the ocean during low tide. Being shallow and partially isolated, the reef flat experiences high variation in temperature throughout the day (3–4°C) and season (6.5°C), with an average water temperature of 20°C in the winter and 28°C in the summer [30]. This experiment took place during Spring (October 7–16) with mean water temperatures of 24°C.

Over the course of 3d, a total of six *Aplysia dactylomela* were found: two Adults and four juveniles (half of which were 2cm long, the other half were 15 cm long) (Fig. 2). All were acclimated for 7d in a large, controlled flow-through tank using water pumped from the reef flat and fed a diet of fresh *Laurencia intricata*.

Algae Treatments

The culturing setup entailed growing *L. intricata* in four different treatments: increased light (L), increased temperature (T), both increased light and temperature (B), and control (C). Twelve 125L shaded flow-through tanks were randomly assigned to treatments, providing three replicates each (Fig. 3). Tanks in the L treatment were illuminated by external LED hood lights from 07:00 to 18:00 each day (Fig. 4A). The T treatment included submersible heaters programmed to raise the temperature of incoming water by 3°C at all times to mimic projected rise of sea surface temperature on the reef flat (Fig. 4B). B tanks were fitted with both hood lights and heaters (Fig. 4C). C tanks were left bare aside from the water inflow tube (Fig. 4D). A shade screen was hung above the tanks to equalize natural light exposure. Two five-gallon buckets of *L. intricata* were collected from the reef flat then divided evenly amongst the 12 treatment tanks and cultured for 5d.

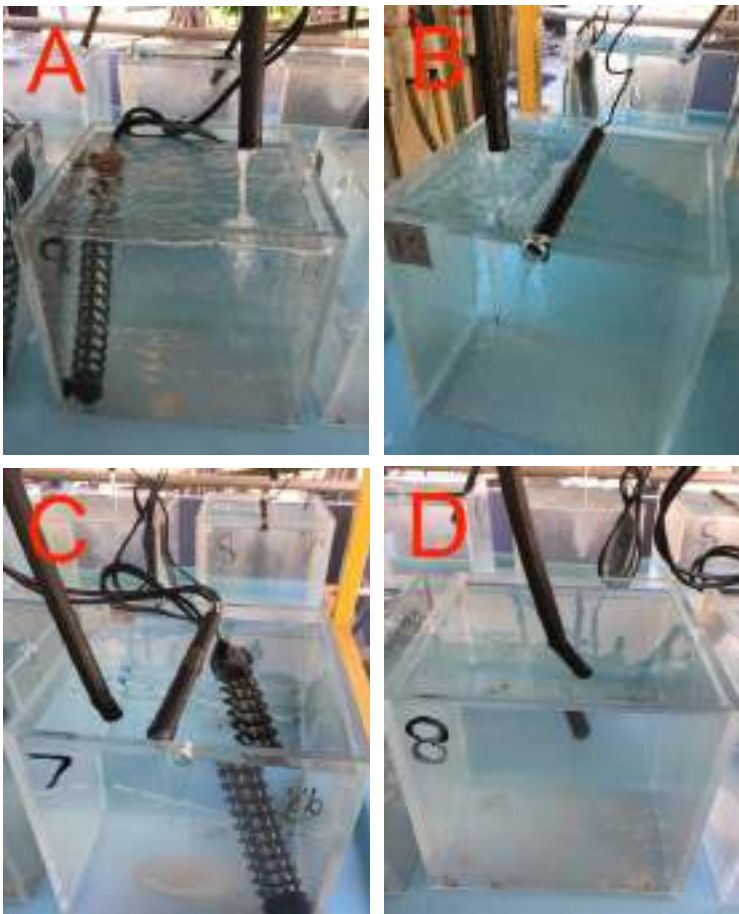


Figure 4. Setup of treatments in which *Laurencia intricata* algae was cultured. All are flow-through tanks using water pumped from the reef flat: **A)** Increased temperature treatment tank with submersible heater. These tanks were kept 3°C above ambient sea surface temperature on the reef flat and were tested daily with a thermometer to verify heater accuracy; **B)** Increased light treatment tank with external LED hood. Lights were positioned directly across the tank at the water surface and were turned on from 07:00 to 18:00 each day. They emitted white light; **C)** Combined light and temperature treatment tank containing both LED hood and submersible heater; **D)** Control treatment tank.

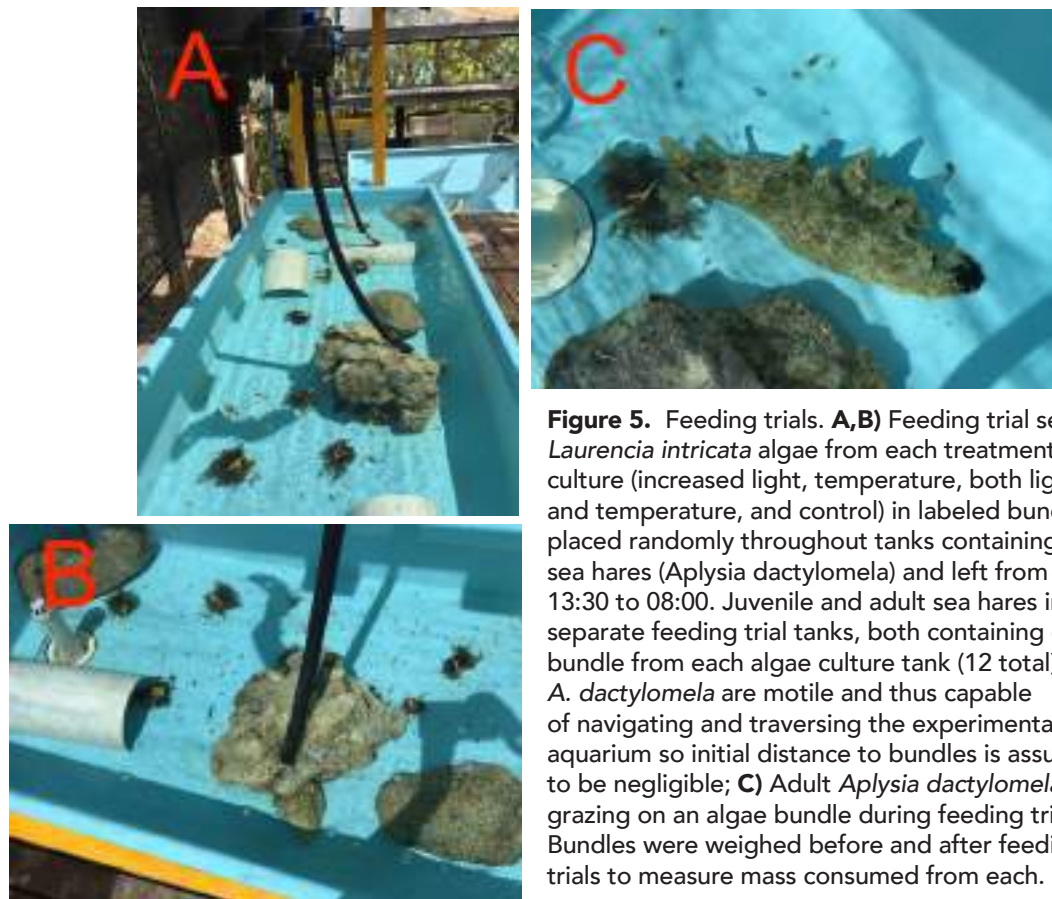


Figure 5. Feeding trials. **A,B)** Feeding trial setup. *Laurencia intricata* algae from each treatment culture (increased light, temperature, both light and temperature, and control) in labeled bundles randomly throughout tanks containing sea hares (*Aplysia dactylomela*) and left from 13:30 to 08:00. Juvenile and adult sea hares in separate feeding trial tanks, both containing one bundle from each algae culture tank (12 total). **A.** *dactylomela* are motile and thus capable of navigating and traversing the experimental aquarium so initial distance to bundles is assumed to be negligible; **C)** Adult *Aplysia dactylomela* grazing on an algae bundle during feeding trials. Bundles were weighed before and after feeding trials to measure mass consumed from each.

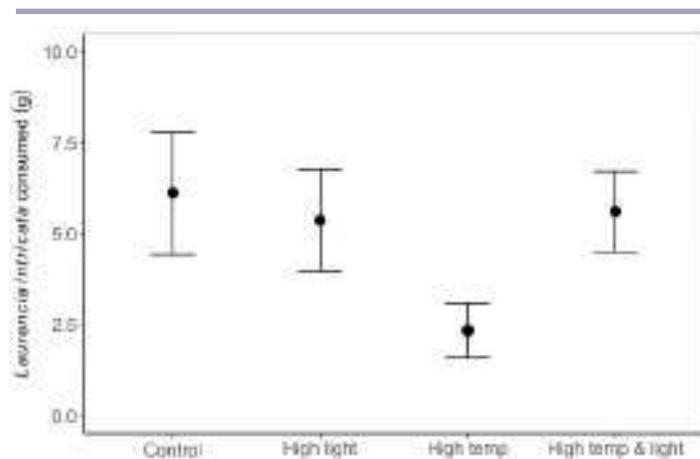


Figure 6. Mean grams of the red algae *Laurencia intricata* consumed per bundle per sea hare from each treatment culture over 18.5h. For control n=6; high light n=6; high temperature n=6, high light and temperature n=4. Bars represent standard error.

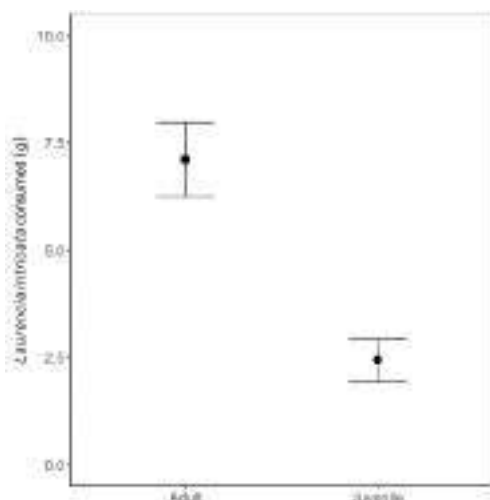


Figure 8. Mean grams of red algae *Laurencia intricata* consumed per bundle per individual by each juvenile (n=4) and adult (n=2) sea hares *Aplysia dactylomela* over a 18.5h period.

Feeding Trials

Algae in each tank was tied into two egg-sized bundles using rubber bands and labeled with treatment and tank number: a total of six replicate bundles per treatment. Bundles were blotted dry with a cloth for 5s then weighed and the mass recorded.

Sea hares were fasted for 12h before the feeding trial. Adults and juveniles were placed into separate tanks with one bundle each from every tank and treatment dispersed randomly throughout to control for differences in initial proximity to sea hares (Fig. 5). Trials began at 13:30 and continued until 08:00 the next morning (18.5h duration) at which time what remained of the bundles was removed, blotted, and weighed again. The difference in mass for each bundle was calculated and associated with the given treatment, tank number, and *A. dactylomela* age. In cases where all algae was consumed, the rubber band and label were weighed. To standardize results between adults and juveniles and achieve per capita values, difference in mass for each bundle was divided by

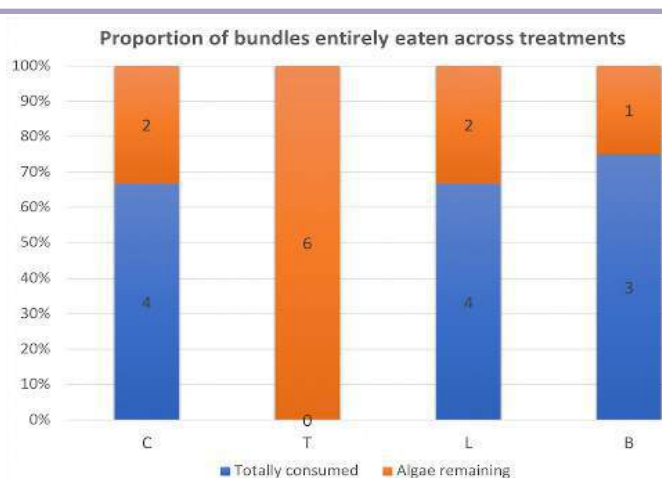


Figure 7. Proportion of red algae *L. intricata* bundles completely consumed by sea hares after 18.5 hours across treatments: Control (C), high temperature (T), high light (L), both high light and temperature (B). Number of bundles in each treatment and category is represented on each bar.

number of sea hares in the tank. The timing of the experiment gave sea hares equal opportunity to feed during different times of both day and night.

Analysis

All data analysis was conducted using R version 3.4.2. and all figures were made through ggplot2. A preliminary QQplot and histogram were used to visually confirm normality of distribution. Data appeared slightly right skewed and were adjusted using a square root transformation (see Appendix). A Levene's test was then run to check the assumption of homogeneity of variance, and a Shapiro test to mathematically verify normality of distribution; both tests passed.

A two-way Analysis of Variance (ANOVA) was used to compare mass of *L. intricata* consumed between algal treatments as well as age groups of *A. dactylomela*. Age and algae treatment were orthogonal fixed factors, while tank number was a random effect of algae treatment. A post-hoc Tukey's test was then used to isolate significant differences between specific treatments.

Results

A. dactylomela ate significantly different quantities of algae across the different treatments ($F_{(3,13)}=15.618$, $p=1.34 \times 10^{-4}$). *A. dactylomela* ate significantly less algae from treatment T than the control or any other treatment ($p < 0.05$), while there was not a significant difference in amount consumed between *L. intricata* treated with C, L, or B (Fig. 6).

Across C, L, and B treatments, approximately 70% of bundles had been completely consumed leaving only the rubber band. However not a single T bundle (0%) was entirely consumed (Fig. 7).

Adults and juveniles ate significantly different amounts: per capita, adult *A. dactylomela* ate significantly more *L. intricata* than juveniles ($F_{(1,13)}=99.865$, $p=1.81 \times 10^{-7}$), with each adult consuming on average 78.37g over the 18.5h period, and juveniles consuming 33.41g each (Fig. 8).

There was no significant difference in the age/treatment

interaction term, meaning adults and juveniles ate the different treatments in similar ways. For full ANOVA and Tukey results see Appendix.

Discussion Findings

Significant results for age likely indicate greater nutrient and metabolic needs for adults than juveniles. Juveniles require energy to grow, and as evidenced by these results, do consume a large amount of algae despite their small size. However, adults were demonstrated to eat over double that amount, as they are larger and likely require more energy to move and produce egg masses. That adults consume higher quantities of algae indicates they have a greater impact individually on algae control than juveniles, which play a smaller but still significant role. However, depending on the mortality rate of *A. dactylomela*, it is possible that there are enough juveniles to together account for the majority of algae consumed. Obtaining measurements for total algae consumed for adults and juveniles also enables future researchers to compare sea hare feeding habits with that of other herbivores to better understand the dominant species in reef macroalgae control.

A non-significant treatment age interaction term reveals that adults and juveniles did not exhibit significantly different preferences between treatments. This knowledge is critical, as it suggests that the changes in sea surface temperature and light exposure predicted to arise will affect the herbivory of all *A. dactylomela* similarly, regardless of age. This is unexpected, as the literature on other species in the *Aplysia* genus suggests juveniles and adults are known to have different preferences among algal species, and consequently, the physical and chemical characteristics of food sources[27]. *A. dactylomela* is unusual in that it specializes on *L. intricata* for its entire life. Thus, in order to gain a more thorough understanding of herbivory preferences across life stages, future experiments would benefit from studying additional sea hare species both within and outside the *Aplysia* genus to determine if this is a trend that holds or a singular characteristic of *A. dactylomela*.

A. dactylomela were shown to eat significantly less *L. intricata* if it had been treated with high temperature. Observations during the culturing process noticed algae in T tanks appeared less pigmented than that of other treatments, and was more fragile to bundle. This could indicate a loss of chlorophyll and nutritional value, making it less appealing to *A. dactylomela*. As algae in T treatments would be equally easy (if not more so) to eat, this aversion is likely caused by a reduction in quality rather than an increase in difficulty [20]. This is consistent with the literature, in which thermal stress on algae has been shown to result in cellular damage and decreased protein, nitrogen, and caloric value[17]. Additionally, algae often responds to this by producing large quantities of heat shock proteins (such as the HSP60 and HSP70 classes) to repair damage[31]. It is possible that these chemical changes could also have an effect on the algae's taste and thus herbivore response.

Interpretations

One interesting result of this study was that despite a highly significant difference for the temperature treatment, there was no significant difference whatsoever for the B treatment of both temperature and light. This could imply that for *L. intricata*, light and temperature are antagonistic with respect to quality as a food source, meaning some feature of increased light is able to mitigate the negative effects of temperature on herbivory preference. This could be a result of an additional 3rd factor such as loading of microbial epiphytes. Low light and higher temperatures in water along with an organic substrate (algae) is known to facilitate bacterial growth[32], which could also make the algae less appetizing. Two 2011 studies [33, 34] on the Australian red alga *Delisea pulchra* found that loss of pigmentation occurs when the it is thermally stressed because warmer temperatures inhibit the alga's ability to produce chemical defenses, leaving it susceptible to bacterial loading and infection that results in loss of chlorophyll and other pigments. This is consistent with our observational evidence of *L. intricata* (also a chemically

defended red alga) lacking coloration and beginning to decay in tanks with increased temperature. Furthermore, the fact that algae in B tanks did not appear any less pigmented, and was shown to be preferred equally to control algae by sea hares, could suggest that increasing light may allow the alga to combat (perhaps through increased photosynthesis and energy production) whatever degradation the increased temperature caused.

Another interpretation is that B and perhaps also L treated *L. intricata* were in fact less nutritious (though still preferable to T), and as a result, *A. dactylomela* (compensated by consuming more of it to meet their needs. This has been shown to occur in several species of amphipods, but as a phenomenon is still largely understudied[35]. If this experiment were repeated, it would be possible to explore this by re-weighing the bundles multiple times during the experiment at shorter intervals to detect a ranking in preference and a succession from the most favored treatments to less desirable ones and gain a better understanding of what is actually occurring. As this experiment is relatively small-scale, and since so little has been studied on the topic, it is difficult to predict how increased temperature and light might impact future reefs, and further research is necessary to develop a more comprehensive understanding of the factors at play. However, the following discussion will outline several possible outcomes supported by these results.

Future Implications

If rising sea surface temperature causes *L. intricata* to lose nutritional value yet continue growing as usual, it could result in less herbivory, as animals unable to meet their energy needs or forced to eat more and travel farther (energetically expensive) to gain sufficient food will likely have reduced growth rate and reproductive capacity. This would result in decreased

“
Rising sea surface temperature could result in less herbivory, as animals unable to meet their energy needs are forced to eat more and travel farther.
”

herbivory and thus allow *L. intricata* to spread and, if other algae and herbivores react similarly, could contribute to a phase shift. Alternatively, if higher sea surface temperatures damage the algae enough to impede its growth and reproduction, as before, this will reduce nutrition available to herbivores, but if unable to grow well even in the absence of predation, algae cover may remain at a low levels. While possibly less harmful to reefs than a phase shift, loss of a key food source can still result in bottom-up effects that reverberate through the food web through changes in trophic interactions. However, one potential positive is that the results for increased light might suggest that herbivory interactions could return to normal during high-light El Niño years. This could be beneficial given that most coral bleaching events occur during El Niño events and thus, when corals are most vulnerable. However, if the constant impact of increased temperature causes enough damage, it may be difficult to recover and re-establish depleted populations over the time scale of an El Niño.

Each of these scenarios have the possibility of contributing to far-reaching changes across the trophic pyramid and pose a threat to reef biodiversity. For this reason, it would be valuable for expand this study to include other species of algae that are dominant and widespread, as well as the most impactful herbivorous reef fish and invertebrates. Additionally, collecting long-term data on the health and fecundity of grazers fed different treatment algae would be beneficial.

Limitations and Considerations

This study was largely exploratory and small-scale in design, thus it is susceptible to all the usual limitations associated with this type of experiment, including small sample size and lack of site-level replication. Additionally, one tank experienced a heater malfunction and was discarded from the experiment so there were only two replicate bundles for B treatment in each age (four total, as opposed to six). Furthermore, only two adult *Aplysia* could be found over the entire three day span in which they were collected. According to Dr. Selina Ward of University of Queensland, who has been studying *A. dactylovela* at Heron Island for over a decade, they are typically highly abundant and finding so few is extremely unusual. This could point to additional factors at play, as whatever is causing such an uncharacteristically low abundance of *A. dactylovela* may also have an impact on the dietary preferences of those found. Some gastropods, such as the sea slug *Placida dendritica* have been shown to exhibit different taste preferences even among individuals of the same population, which is believed to have a genetic component [36]. It is not known if this variability of preference applies to *A. dactylovela* as well, but as only two adults and four juveniles were found, if this were the case the dietary choices of the six individuals studied may not be representative of the species as a whole.

Another consideration worth noting is that while some algae from each bundle was passively lost during the feeding trial (as it is delicate and small pieces likely broke off), this may have occurred disproportionately more for certain treatments. That is, the less healthy and thus more fragile algae may have fallen apart more and incorrectly been counted as eaten.

Acknowledgments

I would like to thank the University of Queensland

and Heron Island Research Station for providing me with the resources and facilities to conduct this research. Additionally, I would like to thank Kristen Brown and Dr. Selina Ward of the University of Queensland for their guidance, knowledge, and support in designing, conducting, and analyzing this experiment. Additionally, I would like to acknowledge the University of Queensland's Jasmine Lee for assistance with statistical tests and Cole Shepherd of Stanford University for help in caring for the sea hares. Lastly, I am thankful for the Bing Overseas Studies Program for enabling me to visit Australia and funding my stay in Heron Island.

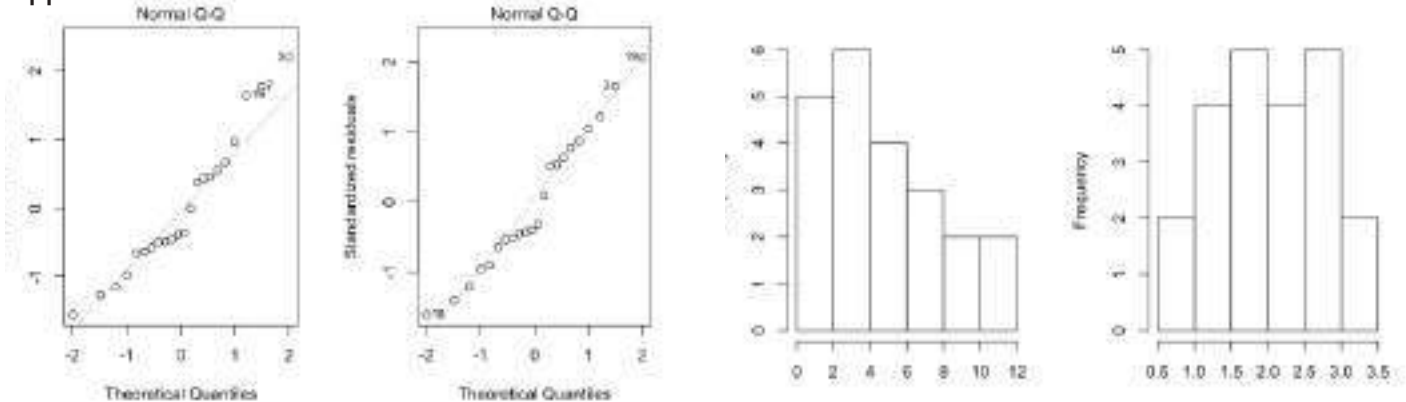
References

- [1] Intergovernmental Panel on Climate Change (IPCC). Climate change 2007: the physical science basis. 2007.
- [2] McMichael AJ, Woodruff RE, Hales S. Climate change and human health: Present and future risks. *Lancet*. 2006.
- [3] Roberts CM, McClean CJ, Veron JEN, Hawkins JP, Allen GR, McAllister DE, et al. Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* (80-). 2002.
- [4] Butler JRA, Wong GY, Metcalfe DJ, Honzák M, Pert PL, Rao N, et al. An analysis of trade-offs between multiple ecosystem services and stakeholders linked to land use and water quality management in the Great Barrier Reef, Australia. *Agric Ecosyst Environ*. 2013.
- [5] Hoegh-Guldberg O, Mumby PJ, Hooten AJ, Steneck RS, Greenfield P, Gomez E, et al. Coral reefs under rapid climate change and ocean acidification. *Science* (New York, N.Y.). 2007.
- [6] Koch M, Bowes G, Ross C, Zhang XH. Climate change and ocean acidification effects on seagrasses and marine macroalgae. *Global Change Biology*. 2013.
- [7] McCook LJ. Macroalgae, nutrients and phase shifts on coral reefs: Scientific issues and management consequences for the Great Barrier Reef. *Coral Reefs*. 1999.
- [8] Done TJ. Phase shifts in coral reef communities and their ecological significance. *Hydrobiologia*. 1992.
- [9] Smith JE, Hunter CL, Smith CM. The effects of top-down versus bottom-up control on benthic coral reef community structure. *Oecologia*. 2010.
- [10] Vergés A, Paul NA, Steinberg PD. Sex and life-history stage alter herbivore responses to a chemically defended red alga. *Ecology*. 2008.
- [11] Bebbington A. Aplysiid species from Eastern Australia with notes on the Pacific Ocean Aplysiomorpha (Gastropoda, Opisthobranchia). *Trans Zool Soc London*. 1977.
- [12] Paul VJ, Pennings SC. Diet-derived chemical defenses in the sea hare *Stylocheilus longicauda* (Quoy et Gaimard 1824). *J Exp Mar Bio Ecol*. 1991.
- [13] Hay ME, Fenical W. Marine Plant-Herbivore Interactions: The Ecology of Chemical Defense. *Annu Rev Ecol Syst*. 1988.
- [14] Barile PJ, Lapointe BE, Capo TR. Dietary nitrogen availability in macroalgae enhances growth of the sea hare *Aplysia californica* (Opisthobranchia: Anaspidea). *J Exp Mar Bio Ecol*. 2004.
- [15] Lough JM. Climate variation and El Niño-Southern Oscillation events on the Great Barrier Reef: 1958 to 1987. *Coral Reefs*. 1994.
- [16] Josselyn MN. Seasonal changes in the distribution and growth of *Laurencia poitei* (rhodophyceae, ceramiales) in a subtropical lagoon. *Aquat Bot*. 1977.

- [17] Mendes LF, Vale LAS, Martins AP, Yokoya NS, Marinho-Soriano E, Colepico P. Influence of temperature, light and nutrients on the growth rates of the macroalga *Gracilaria domingensis* in synthetic seawater using experimental design. *J Appl Phycol*. 2012.
- [18] Rosen G, Langdon CJ, Evans F. The nutritional value of *Palmaria mollis* cultured under different light intensities and water exchange rates for juvenile red abalone *Haliotis rufescens*. *Aquaculture*. 2000.
- [19] Steneck RS, Watling L. Feeding capabilities and limitation of herbivorous molluscs: A functional group approach. *Mar Biol*. 1982.
- [20] Nicotri ME. Factors involved in herbivore food preference. *J Exp Mar Bio Ecol*. 1980.
- [21] Lapointe BE, Barile PJ, Yentsch CS, Littler MM, Littler DS, Kakuk B. The relative importance of nutrient enrichment and herbivory on macroalgal communities near Norman's Pond Cay, Exumas Cays, Bahamas: A "natural" enrichment experiment. *J Exp Mar Bio Ecol*. 2004.
- [22] Clarke CL. The Ecological Role of Sea Hares (Opisthobranchia: Anaspidea) within Tropical Intertidal Habitats. James Cook Univ (Masters Thesis). 2006.
- [23] Pennings SC, Paul VJ. Effect of plant toughness, calcification, and chemistry on herbivory by *Dolabella auricularia*. *Ecology*. 1992.
- [24] Pennings SC, Nadeau MT, Paul VJ. Selectivity and growth of the generalist herbivore *Dolabella auricularia* feeding upon complementary resources. *Ecology*. 1993.
- [25] Geange SW, Stier AC. Charismatic microfauna alter cyanobacterial production through a trophic cascade. *Coral Reefs*. 2010.
- [26] Gerdes R, Fieber L a. Life history and aging of captive-reared California sea hares (*Aplysia californica*). *J Am Assoc Lab Anim Sci*. 2006.
- [27] Pennings SC. Size-related shifts in herbivory: specialization in

- the sea hare *Aplysia californica* Cooper. *J Exp Mar Bio Ecol*. 1990.
- [28] Poore AGB, Hill NA. Sources of variation in herbivore preference: Among-individual and past diet effects on amphipod host choice. *Mar Biol*. 2006.
- [29] Burkepille DE, Hay ME. Herbivore species richness and feeding complementarity affect community structure and function on a coral reef. *Proc Natl Acad Sci*. 2008.
- [30] Georgiou L, Falter J, Trotter J, Kline DI, Holcomb M, Dove SG, et al. pH homeostasis during coral calcification in a free ocean CO₂ enrichment (FOCE) experiment, Heron Island reef flat, Great Barrier Reef. *Proc Natl Acad Sci*. 2015.
- [31] Vayda ME, Yuan ML. The heat shock response of an antarctic alga is evident at 5°C. *Plant Mol Biol*. 1994.
- [32] Pomeroy LR, Wiebe WJ. Temperature and substrates as interactive limiting factors for marine heterotrophic bacteria. *Aquatic Microbial Ecology*. 2001.
- [33] Case RJ, Longford SR, Campbell AH, Low A, Tujula N, Steinberg PD, et al. Temperature induced bacterial virulence and bleaching disease in a chemically defended marine macroalga. *Environ Microbiol*. 2011.
- [34] Campbell AH, Harder T, Nielsen S, Kjelleberg S, Steinberg PD. Climate change and disease: Bleaching of a chemically defended seaweed. *Glob Chang Biol*. 2011.
- [35] Cruz-Rivera E, Hay ME. Can quantity replace quality? Food choice, compensatory feeding, and fitness of marine mesograzers. *Ecology*. 2000.
- [36] Trowbridge CD. Diet specialization limits herbivorous sea slug's capacity to switch among food species. *Ecology*. 1991.

Appendix



Appendix I. A) Histograms and **B)** QQplots of the data before (left) and after (right) implementing the square root transformation.

DF	Sum Sq	Mean Sq	F value	Pr(>F)	
Age	1	7.737	7.737	99.865	1.81e-07 ***
Algae_Treatment	3	3.638	1.210	15.618	0.000134 ***
Age:Algae_Treatment	3	0.176	0.059	0.757	0.537670
Residuals	13	1.007	0.077		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

diff	lwr	upr	p adj	
C-B	-0.007981336	-0.7658166	0.7498540	0.9999892
L-B	-0.120177513	-0.8780128	0.6376578	0.9663406
T-B	-0.921935302	-1.6797706	-0.1641000	0.0153882 *
L-C	-0.112196177	-0.7908247	0.5656323	0.9620415
T-C	-0.913953966	-1.5917825	-0.2361255	0.0074320 **
T-L	-0.801757789	-1.4795863	-0.1239293	0.0185271 *

Appendix II. Full results from the 2-way ANOVA with significant values highlighted in yellow.

diff	lwr	upr	p adj	
Juvenile-Adult	-1.343031	-1.512436	-0.7736256	1.12e-05 ***

Appendix III. Results from TukeyHSD test performed on **A)** treatment (top), and **B)** age (bottom). Significant results highlighted in yellow.



STANFORD UNDERGRADUATE
RESEARCH JOURNAL

VOLUME 19 | SPRING 2019

Stanford, CA 94305

surj.stanford.edu