SIGNAL TO NOISE: INTRA-ACTIVE ENTANGLEMENTS IN AN INTERDISCIPLINARY COURSE ON DATA AND STORYTELLING

By

Lee Melvin M. Peralta

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Curriculum, Instruction, and Teacher Education—Doctor of Philosophy

2024

ABSTRACT

In this dissertation, I engage in three analytic cuts to think about/with a relational ontological orientation to data and data literacies/science education. The analysis focuses on the following question: What possibilities for teaching and learning about data are made possible when we attune to the relational, noisy, liminal, and material dimensions of data and the connections between data and broader issues of power and ethics? This study is situated in an interdisciplinary course on data and data storytelling at a large public university in the U.S. Midwest that I and another colleague designed and taught in the 2022-2023 academic school year. I organize my findings and discussions along three chapters in this dissertation. Chapter 2 is a theoretical examination of norms and values about data suggested in certain data science education reform efforts in the U.S. and a reconsideration of new possibilities for teaching and learning about data informed by relational ontologies and philosophical theories about signal and noise. Chapter 3 is an empirical piece that shares vignettes of two students' engagements with data physicalizations as part of a data postcard activity and year-long survey-based research project. Chapter 4 examines how opportunities for critical, creative, and interdisciplinary engagements with data throughout the data storytelling course shaped how students made sense of data and processes of data generation, analysis, and communication. I co-authored the piece alongside my co-instructor and five of the students from the data storytelling course. Overall, this dissertation offers a unique approach of attuning to and elevating the concept of noise as a potentially generative concept for data literacies/science education. It raises important questions about the role of material agency, ethics and response-ability, ambiguity and improvisation, storytelling, and interdisciplinarity in connection with efforts to teach and learn about data in critical, creative, and relational ways.

To Vida, Leo, Riza, David, Carisa, and Jojo

ACKNOWLEDGEMENTS

First, I would like to thank my co-instructor and the students of the data storytelling course. My dissertation would not have been possible without them. This dissertation was also motivated and shaped by my former experiences as a mathematics teacher at Advanced Math and Science II in the Bronx and Louis Pasteur Middle School in Queens. For this reason, I would also like to thank my former students, my former co-workers, and my friends and former colleagues from my time as a fellow in Math for America, which enabled me to transition into teaching and which provided a strong foundation for the rest of my career in education.

Words cannot express my gratitude to my advisor and chair of my dissertation committee, Dr. Beth Herbel-Eisenmann. She supported me throughout my entire time at Michigan State University, listening to and providing feedback and guidance on my emerging scholarly interests examining the historical, sociopolitical, and aesthetic dimensions of mathematics and data science education. She also served as an invaluable mentor and colleague on two research projects, namely one involving a multi-year collaboration with scholars across the United States and Israel working to develop critical theoretical frameworks and commentaries on data science education and another involving a multi-year partnership with youth and youth leaders in a local subsidized housing community. I am also grateful for the support from my dissertation committee. Dr. Joanne Marciano, through our work in a youth participatory action research project alongside Dr. Herbel-Eisenmann, taught me the value of building and nurturing relationships as part of the work of teaching, research, and being a scholar. Dr. Higinio Dominguez introduced me to the world of relational ontologies such as those from feminist new materialisms and various traditions of Indigenous thought. He has had a profound influence on the direction of my scholarly interests and of this dissertation study. Dr.

Sandro Barros consistently showed genuine curiosity and passion for my ideas about data and introduced me to the intellectual rabbit hole that is object-oriented ontologies, alien phenomenology, and speculative realism. Dr. Nathalie Sinclair generously gave her time and energy toward providing incredible guidance and feedback as my external committee member. I have long admired her work and continue to grow as a result of her scholarship about aesthetics and time and their connection to education.

I am deeply indebted to faculty and staff from outside of my committee. I would like to extend my deepest gratitude to Dr. Louise Jezierski, without whom my dissertation would not have been possible. She not only served as an incredible co-instructor and colleague but also as a mentor and friend, particularly through the Interdisciplinary Inquiry and Teaching Fellowship Program at MSU. I also thank Dr. David Stroupe for his course on sociocultural learning theories; Dr. Lynn Fendler for her course on philosophy of education; Dr. Tonya Bartell for her course on mathematical ways of knowing; Dr. Dorinda Carter Andrews for her course on critical race theory; Drs. Aman Yadav and Michael Lachney for their course on race, education, and technology; and Dr. Sandro Barros for his course on humanities-oriented research. These courses had a notably influential impact on my development as a scholar and educator. I would also like to thank Dr. Alicia Alonzo, Dr. Sandra Crespo, Freda Cruel, Lisa Keller, Terry Edwards, Lori Jean Nichols, Taylor Gallagher, and Dawnmarie Yates who, in their leadership, administrative, and support capacities, made my time as well as those of everyone else at MSU possible. I thank Dr. Anne-Lise Halverson for her support and for convincing me to attend MSU. I thank Dr. Anna Pegler-Gordon for her leadership in the Interdisciplinary Inquiry and Teaching Fellowship during the COVID-19 lockdowns. Outside of MSU, I extend my gratitude to my mathematics and data science education collaborators Drs. Laurie Rubel, Vivian Lim, Jennifer Kahn, Travis

Weiland, and Shiyan Jiang (and Beth Herbel-Eisenmann), along with Drs. Carlos Lopez Leiva and Glady Kraus (and Higinio Dominguez). I thank Dr. Cheryl Matias for giving me a chance on my first solo authored book chapter publication. I thank Dr. Sunita Vatuk for her mentorship and Dr. Despina Stylianou for helping me become a teacher and for continuing to believe in me throughout my work with The City College of New York.

The community of graduate students, colleagues, and friends at MSU and beyond has been one of the most important aspects of my time as an emerging scholar. The reasons for my gratitude, their contributions to my life, and the ways our lives and experiences intersect and influence one another are so rich and complex that I will not even attempt to spell them out here. Instead, the best that I can do is honor them here, so I would like to shout out: Sofía Abreu, Selin Akgün, Joaquín Argüello, Joel Berends, Alecia Beymer, Julianna Blain, Sunghwan Byun, Grace Chen, Kyle Chong, Darshana Devarajan, Benjamin Dickman, Anthony Dickson, Christopher Dubbs, Kasun Gajasinghemliyanage, Nicolei Gupit, Regina Gong, Ezra Gouvea, Hannah Grisham, Reyila Hadeer, Patti Hamerski, Jermaine Howell, Benjamin Iremonger, Brent Jackson, Lauren Elizabeth Johnson, Ahmad Wachidul Kohar, Ji Soo Lee, Rileigh Luczak, Marisol Masso, Bri Markoff, Alexander Moore, Helena Omaña Zapata, Sheila Orr, Molade Osibodu, Byungeun Pak, Arienne Patano, Jess Reed, Sol Rheem, Rachel Roca, Miguel Rodriguez, Ishan Santra, Eric Shepherd, John Sweeney, Brady Tyburski, Vivek Vellanki, Katie Westby, Reneé Wilmot, Al Wood, and Sabrina Zarza.

I am grateful to my family for their support. This dissertation would not be possible without my parents, Vida and Leo Peralta, who taught me and my siblings the value of education. I will forever be grateful to my siblings, Riza and David Peralta, who make Maryland feel like my true home. I welcome my new sister-in-law Carisa Vercaza. I shout out our aunties,

uncles, and cousins: Auntie Loida, Uncle Art, Lala, Michelle, Roda, Diana, Auntie Edna, Rachel, Auntie Josie, Uncle Romulo, Jomie, RJ, Jojo, Uncle Eddie, Joyce, Lutz, and Sheila, along with my Solon family not related by blood: Jennifer Wilmore, Tara Gonsalves, Michael Lopez, and Caroline Honsa, and all of their partners and children. Last but not least, I send my love to the cats: Jojo, Phoebe, Poptart, and Apple. They are just as much a part of the family as anyone else  .

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

In connection with an interdisciplinary course on data and storytelling offered at a large university in the U.S. Midwest, each of the nine undergraduate student participants from the course shared their developing ideas about data through several individual and focus group interviews:

I feel like with every story now, there's a meta story of who wrote it and how. That impacts the story. And to me, before data was just hard facts, but now there's so much more behind it that I didn't think about before. So it's almost like for every story that's told, there are stories outside of the story to think about. - Adele[1]

I consider data a concept. Data doesn't really exist. It's sort of our interpretation of something in reality. And it's never going to be absolute, like sure you can, you can interview everybody at [university] and still would not get a truthful interpretation because you oozed your bias all over the questions and all over how you collected it and you did it through a computer or through a tablet or anything like that and then after you've got all that oozed over collection, and then you try and put it in categories, which data resists because it's automatically not a thing that exists. So you're putting a concept into a bunch of other concepts. And then after that, you're trying to turn it into a physical entity, and in the creation of that, through the materials you use, through the people who use it, through the time you do it, you are also oozing bias in your own interpretation onto it. So you've got an interpretation of interpretations of an interpretation of an interpretation of data. So it's barely the same thing. - Bobble

It has become more of a creative outlet than I had perceived before. Being able to do my own visualization and being able to do a little drawing and color things and make it look fun and have something creative. I mean, there's a million ways that you could represent data. I would think of it more of a square thing, and now I think of it more as like a circular thing, if that makes sense. Just, it's like more fluid than it was before. Like it's a lot less rigid. - Constance

I just thought data was in the details. If you gave me data visualization, I would've picked out the data in it, but I would've thought of like individual data points and the visualization. But now I see it as like a whole picture and that there's like more to data than the details. - Daisy

What I have learned is that, similar to writing, data storytelling is a process. It's not a solution. There's not one correct answer, especially with the data postcard where we have been asked to recreate that same data in different ways... And especially with the Legos, it really makes you dive into every detail of data that you wouldn't necessarily have in order to create the story that you want to tell. - Dawn

---

[1] All names are pseudonyms.

My understanding of the term has changed because now I catch myself sitting here and I'm like, oh, data. I have to think about it, and I'm realizing I used to consider data a straight number that usually had a decimal point and was really specific. And then, you know how you talked about the clock? That sits in my head all the time. The clock example: How data can be an actual thing or data is subjective and all of that. So my understanding has changed, and I honestly right now could not define the term data because it lowkey still confuses me. - Eve

You do have to make choices. We saw that with the data story visualizations for our screen time assignment. When you choose to represent things in different ways, you lose, in almost all cases, you lose something. It takes a very skilled person to not lose any data when they're visualizing it in different ways. But those are choices that you have to make, and sort of like the ducks, you can't control all the ducks. So when you display things in different ways, you invoke different emotions, you invoke different reactions. - Frodo

My positionality has definitely influenced the way I look at data, especially when it refers to income groups in health. I have always been interested in pursuing medicine because it's a dream of mine to be able to help people in the future. I suppose one way my positionality has been shaped is that now when I look at data relating to income inequality and COVID-19 within different income groups from my Data Exploration 1, I realize that I've been categorizing myself incorrectly and I'm actually from an upper middle class family. Though this doesn't seem like much of a difference, I definitely think that it's an important distinction to recognize. - Saba

I was never really a big fan of data in and of itself, but the way that you guys had described it being stories, me as an English major, absolutely that kind of caught my eye because I was like, okay, now that sounds interesting. Because you're taking something that on the surface sounds boring, like blah data, whatever, boring, you know? And then you're saying, hey, there's more to that. There's stories behind this. And that's probably my favorite thing about the course is learning that there are stories behind data. And I loved that because the English major in me thrived because I love writing, hence the creative writing focus. I absolutely love anything that has to do with that. I'm big for a story. - Sasha

The responses by the student participants suggest a range of developing ideas and questions about data that emerged within the data and storytelling course. Adele, Daisy, and Dawn, for instance, hint toward broadening their conceptions toward data, where data are no longer “individual data points” but rather stories where “for every story that's told, there are stories outside of the story to think about.” The idea of telling stories through data, coupled with the idea of data as stories, resonated with students such as Sasha, whose background in creative writing led her to state that although data “on the surface sounds boring, like blah data,” “there’s more to that” and “there are stories behind data.” For Eve, although the course on data and

storytelling changed her mind about data "as a straight number," the course also left an open-endedness to data that left her "confused" about its definition. For Frodo, the idea of data and storytelling—which the students developed while engaging in activities involving data postcards, screen time assignments, data story explorations, and a survey-based research project—highlighted the role of choices in visualizing and representing data. In particular, her statement of "losing something" while making data-related choices echoes Barad's (2007) notion of enacting boundaries through "agential cuts," an idea that plays a central role throughout this dissertation study. Bobble's statement that data are "an interpretation of an interpretation," which is shaped not only by the discursive act of interpreting data but also by the physicalization of data involving the use of time and materials, further resonates with Barad's agential realism. Finally, for Saba, data are shaped by one's positionality, but the inverse is also true: engaging with data can change one's understanding of oneself.

This range of student responses provokes the following question: What possibilities for teaching and learning about data were made possible by the course, and what possibilities were foreclosed? One approach to studying this question would be to closely examine the words of the student participants and, using coding systems such as those detailed in Saldaña (2021), look for patterns across the available data. Following critiques of coding as a method of inquiry (e.g., Augustine, 2014), along with interest among some scholars to interrogate widespread conceptualizations of data and data science (e.g., McQuillan, 2018), however, I use this dissertation study as an opportunity to engage in data science education research in ways that depart from approaches that rely too heavily on coding and that do not ask critical questions about the nature of data, data science, and data science education. Therefore, this dissertation serves the broader goal of seeking out new possibilities for data literacy and data science

education that embraces alternative forms of research and attends to epistemological and ontological orientations to data that emerge within classroom environments conceived in both discursive and material terms.

Study Context

This dissertation study takes place during a time of increasing interest in data literacy and data science education across K-16 settings. Because of the increasing importance of data in society, efforts and programs such as YouCubed at Stanford University, the Data Science Academy at North Carolina State University, and the Introduction to Data Science (IDS) Curriculum at the University of California Los Angeles are producing curriculum, course offerings, teaching materials, professional development opportunities, and research, all in an effort to integrate and foreground data literacy and data science within school curricula. The term data is broad with no consensus on its meaning (Mertala, 2020). For purposes of this study, I use the term data in the sense that many of these organizations appear to use the term: as structured or unstructured collections of numbers, pictures, words, sounds, and other quantifiable, recorded phenomena situated within particular contexts (see Bargagliotti et al., 2020). Tools and concepts commonly associated with this conceptualization of data include, but are not limited to, descriptive statistics and statistical inference; computational statistics; data collection, storage, and cleaning; and representational techniques such as visualization, sonification, and physicalization.

Data science is an interdisciplinary endeavor. Klein (2010) states that interdisciplinarity is an act that involves integrating, interacting, linking, focusing, and blending different fields together. In contrast, multidisciplinarity is a matter of juxtaposing disciplines. In both interdisciplinarity and multidisciplinarity, “disciplines remain separate, disciplinary elements

retain their original identity, and the existing structure of knowledge is not questioned" (p. 17). In contrast, Klein (2010) associates transdisciplinarity with transcending, transgressing, and transforming disciplines and disciplinary boundaries. Following Klein's observation that "many so-called 'interdisciplinary' curricula are actually a multidisciplinary assemblage of disciplinary courses" (p. 17), it possible to argue that many current trends in data literacy and data science education are in fact multidisciplinary or transdisciplinary efforts rather than merely interdisciplinary ones insofar as conversations about data science education inherently involve reconsidering the boundaries of traditional school subjects such as mathematics, statistics, science, the social sciences, the arts, and the humanities. In this study, I use the term interdisciplinary to describe the data storytelling course because it was the term used by me and my co-instructor when we first designed the course and because I use the term interdisciplinary broadly to encompass the possibility of multidisciplinary and transdisciplinary enactments within data science education.

Although this study's central focus is not on the distinction between these terms, they are worth mentioning because the question of what counts and does not count as legitimate disciplinary activity is an important theme of this study and an underlying tension within broader conversations about data literacy and data science education. Despite the inter/multi/trans-disciplinary nature of data literacy and data science, many of the efforts being made in data science education originate in or are connected to statistics and mathematics education (e.g., Bargagliotti et al., 2020; LaMar & Boaler, 2021). Moreover, the learning sciences, particularly in connection with the 2020 special issue on data science in the *Journal of the Learning Sciences*, has been especially influential on my personal learning journey in relation to data literacy and data science education.

Background and Teaching Philosophy

Because this dissertation study is grounded in an undergraduate course in which I took the lead designing and teaching, a discussion of my teaching philosophy is relevant. Before entering my doctoral program, I taught mathematics in a charter high school and then public middle school in New York City. Through my teacher education program at The City College of New York, along with professional development from Math for America, I developed a commitment to student-centered, inquiry-based instruction. This commitment began as early as my first year as a teacher in a relatively new charter high school in the Bronx, where the lack of a prescribed or prepared curriculum enabled me to experiment with the design and implementation of inquiry-based mathematics tasks that sought to go beyond the development of rote skills and procedural fluency. Despite the trial and error and setbacks that I experienced as a first-year teacher, the students were generally receptive to these inquiry-based approaches to teaching. I believe that some of my success can be attributed to my strong interest in mathematics, which began when I was an undergraduate student studying mathematics. There, I developed a passion for mathematical thinking, going so far as to adopt a Platonic worldview of mathematics that saw the discipline as intellectually pure and divorced from the messy contingency of physical reality. Inspired by the writings of scholars such as Lockhart (2009), who suggest the need for exposing students to the beauty of mathematics as early as possible, I sought to inspire in my students an appreciation for mathematics as a field defined by mental intuition, problem solving, and elegance. In retrospect, while my approach and philosophy to mathematics and mathematics teaching may have resonated with many students, it is likely that it did not work well for others.

Despite my perceived initial successes in the Bronx, administrative changes occurred later in the year that required me to adopt my teaching style to a more traditional skill-based

approach to mathematics teaching. For this reason, I transitioned to working in a public middle school in Queens. Nestled in the eastern edge of Queens in an area of New York City adjacent to Long Island, the school had a majority APIDA (Asian American Pacific Islander Desi American) student population, with backgrounds across a wide socioeconomic spectrum. Many families had expectations for and resources to enable their children to attend some of the most prestigious public high schools in New York City and, as such, a part of the school culture emphasized helping students receive high grades and standardized testing scores. In my first year at this new school, I continued to focus on how individual students constructed mathematics knowledge and how best I could support them in this endeavor. I also began to see how my identity as a Filipino American had a positive impact on some of my students, who were attending a school where the majority of teachers were white. I started to recognize patterns of discourse around schooling and mathematics that resonated with my own upbringing while, at the same time, witnessing that a disproportionately low number of Black and Brown students were enrolled in the accelerated courses offered at the heavily tracked school. These experiences were a turning point for me because they led me to realize that teaching and learning are sociopolitical activities shaped by a myriad forces within and outside the classroom, including student and teacher discourse; educational policy enactments; and historical and contemporary patterns of race, gender, class, and ability-based privilege and oppression, among other forms of social hierarchy and violence.

I entered graduate school to better understand classroom discourse and identity development within mathematics classrooms. I engaged with scholarship, explored ideas, and collaborated with mathematics teacher educators and pre-service and in-service mathematics teachers in ways that would not have been possible when I was a full-time teacher. These interactions and opportunities led me to shift my teaching and research focus toward an area of

activity that I had previously been unaware of. This area of activity concerns the idea that it is possible and in many cases desirable to question the nature of mathematics itself, especially with respect to the kind of mathematics typically taught in classrooms. This was another significant turning point for me as an educator because I had previously held a fixed view of what mathematics was or could be. Although I had prided myself in going beyond the procedurally-focused approach to mathematics that too often is promoted in schools, my vision of mathematics as pure, abstract, and universal was nevertheless limiting. In this limited view, mathematics was a beautiful albeit fixed discipline, and my role as a teacher was to employ inquiry-based teaching practices to help move students towards this fixed point. Although I gave serious thought during my time as a teacher to how best to serve my students, particularly those from historically marginalized backgrounds, and help them succeed at mathematics, I took for granted the possibility of questioning mathematics itself.

Through the opportunities afforded to me in graduate school, which included coursework on mathematical ways of knowing and research assistantships that pushed my thinking about the nature of mathematics, I have learned that Platonism is only one among many ways of conceptualizing mathematics. Equally important has been the realization that mathematics, as a discipline and activity with unsettled boundaries, plays an active and not passive role in mathematics classrooms. Chen (2023) states, "Mathematics and mathematics education are also Others in relation, in the form of 'agentic material-discursive practice that… work on and with human and other-than-human existences' (Boylan, 2017, p. 5), rather than merely a context in which teacher-student relations occur" (p. 285). Her words are compelling to me for two reasons. First, they provide language to disavow the notion that mathematics is an already well-defined concept and that there is no need to attune to further possibilities for what it could be beyond

what is commonly considered adequate mathematical knowledge for teaching. Instead, consistent with my growing appreciation for scholarly traditions that frame the world in terms of dynamic, contingent, and mutually co-constitutive relations, mathematics is an agentic participant—or "Other in relation"—in a mathematics classroom, alongside the people who inhabit such classroom. Second, Chen's words reinforce what many people and communities have already stated or suggested about mathematics: that it is not merely an abstract pursuit divorced from physical reality but rather a phenomenon that "work on and with human and other-than-human existences." For these reasons, I have used my time as an emerging scholar to seek out and consider otherwise possibilities for mathematics, turning first toward ethnomathematics (D'Ambrosio, 1985; Barton, 1996) to understand non-dominant traditions of mathematical practice, and then toward approaches that rethink mathematics as a political, embodied, and aesthetic phenomenon (e.g., de Freitas & Sinclair, 2014; Dominguez et al., 2023; Palmer, 2011).

Expanding my thinking about mathematics has since resulted in a positive feedback loop that has also led me to broaden my thinking of teaching and learning generally. Bozalek and Zembylas (2017), for instance, use the term response-able pedagogies to refer to ethico-political educational practices that "incorporate a relational ontology into teaching and learning activities and thus extend their transformative potential" (p. 64) while also acknowledging power relations, materiality, and entanglement in teaching and learning. In a relational ontology, students are not mere cognitive beings who construct knowledge within mental schema but rather are part of a meshwork (Ingold, 2011) of relational processes "through which social, political, and material entanglements in education (i.e. students, facilitators, discourses, texts, performances, drawings, face-to-face and online comings-together) are rendered capable through each other to bring about social transformation" (p. 64). These ideas shaped my approach to instructional design and

teaching throughout the course that grounds this dissertation study. In my role as an instructor, I engaged in the everyday work of teaching—which included planning and facilitating lessons, getting to know students, adjusting the course based on their interests and class contributions, assessing student learning, providing feedback, and collaborating with my co-instructor—but I also came to the course with intentions informed by my developing beliefs about teaching and learning. My first intention was to attune to the classroom as a space of response-ability: how was I responding to my students, my co-instructor, and the material environment of the classroom, and how was the course rendering all of us capable of response? My second intention was not to take for granted the contingent nature and agentic capacity of data. How are data working as an "Other in relation" within the classroom rather than "merely a context in which teacher-student relations occur"? Both of these questions are focal points for each of the three core chapters of this dissertation, for which I provide an overview in the next section.

Before proceeding with the chapter overviews, it is worth addressing a final question: why am I seeking otherwise possibilities for teaching and learning about data in the first place? I came to this dissertation with a background in mathematics education, and yet this study is not situated in the field of mathematics education in a traditional sense. Because of circumstances that led me to pursue a master's in statistics while pursuing my doctorate in education, my teaching and research interests during graduate school gradually shifted toward data literacy and data science education. Many efforts to expand data literacy and data science are connected to mathematics education. Therefore, I continue to see my teaching and research activities as relevant to the mathematics education community. This is especially true because of efforts underway at the time of this writing in certain states such as California to make data science a greater part of K-12 mathematics curricula.

Moreover, data literacy and data science education is a particularly ripe area for seeking out new possibilities for teaching and learning about data, including new possibilities for conceptualizing the notion of data itself. This is the case for several reasons. Data literacy and data science education are nascent fields, and therefore expectations for a "typical" data literacy or data science classroom have not yet been widely established. Further, the fact that data is an interdisciplinary topic that lacks consensus on its meaning is both a challenge for the field of data literacy and data science education but also an opportunity to take risks and propose novel ideas about data that are not limited to those perspectives foregrounded by mathematics, statistics, and computer science. Finally, as will be discussed throughout this dissertation study, scholars have critiqued conventional understandings of data as a tool that enables its users to understand the world from a detached and disinterested perspective (D'Ignazio & Klein, 2020), while others have pointed out historical and contemporary misuses of data (e.g., Chun, 2021; Eubanks, 2018; Noble, 2018; O'Neil, 2016). These issues make reimagining data literacy and data science within classrooms—both in terms of instructional design and pedagogy and in terms of interrogating the nature of data itself—not only appropriate but urgent.

Chapter Overviews

The three core chapters of this study address the question of investigating new possibilities for teaching and learning about data in distinct ways. Although at times they reference one another, they are intended as independent manuscripts that I will submit to academic journals. Each chapter speaks to a particular audience, which I specify in this section.

Chapter 2 is a theoretical examination of the norms and values about data suggested in certain data science education reform efforts. Particular attention is paid to the prominent U.S. policy document about statistics education called the *PreK-12 Guidelines for Assessment and*

Instruction in Statistics Education II (GAISE II, Bargagliotti et al., 2020), U.S.-based initiatives such as Data Science for Everyone, and a report about data science education for undergraduate students published by the National Academies of Sciences, Engineering, and Medicine (NASM, 2018). As a part of my analysis, I read Barad's idea of agential realism through the concept of noise in order to surface the interplay between signal and noise in conversations about data and to provoke new questions for data science education. The purpose of this theoretical discussion is to consider broad epistemological and ontological questions about data, including how data shape and are shaped by human and more-than-human activity and how racialized, gendered, classed, and abled bodies materialize and are materialized through data. In the style of a traditional dissertation, this opening core chapter provides conceptual framing relevant to the remainder of the dissertation study. I prepared this chapter to submit to the journal *Cultural Studies ↔ Critical Methodologies*.

Chapter 3 concerns the possibilities about teaching and learning about data that were enabled and foreclosed in the data storytelling course on which this study is based. The chapter focuses on the activities of two students, Bobble and Saba, as they engaged in a data postcard activity and created a data physicalization made out of yarn as part of a year-long survey-based research project. My analysis employs an intra-active analysis that attends to how human and more-than-human bodies and affective intensities within the course enacted boundaries and exclusions of concept formation that reproduced, at times, and disrupted, at other times, traditional orientations toward data. Findings suggest that when researchers and practitioners seek to make sense of the rich landscape of data education reform, they must confront deep onto-epistemological questions about the nature of data, namely questions about materiality, time and space, and power. Instead of outright accepting or rejecting any one particular approach to data

education, researchers and practitioners should attune to the meshworks of understandings and practices that emerge within data-intensive learning environments and to carefully consider which onto-epistemological assumptions about data to promote and which to challenge in their classrooms. I plan to submit this chapter to *Educational Studies in Mathematics*.

Chapter 4 is the result of an authoring collaboration among me, my co-instructor, and five students from the course who agreed to become co-authors of the manuscript on which Chapter 4 is based. This chapter concerns the data storytelling course and examines students' evolving understandings of data based on their interdisciplinary backgrounds, the interdisciplinary nature of the course, and the novel approach of the course to teach about data in critical and creative ways. The manuscript on which Chapter 4 is based began when Dr. Jezierski and I proposed two research questions for the chapter. These research questions were based on conversations that we had held for the past two years and that were developed during a pilot study of this project. I drafted some initial language for the manuscript, including a literature review about data literacy and data science education that would help contextualize the remainder of the paper. Dr. Jezierski and I then invited students to join the paper as co-authors based on the tentative language in the manuscript. After a subset of five (out of nine) of the student participants agreed to serve as co-authors, we held meetings to discuss ideas for the paper and how responsibilities would be divided among the paper's authors. Dr. Jezierski and I also held individual meetings with students to discuss their ideas and how they wanted to contribute to the paper. This structure enabled each student to have a prominent voice in the paper and allowed Dr. Jezierski and I to provide informal mentoring for the students, many of whom had never undergone the process of co-authoring a manuscript for publication. During these conversations, the paper evolved and changed significantly. For instance, one of the two research questions was removed entirely

because the student co-authors were most engaged with what would become the sole research question of the paper. These events highlighted the powerful impact of including students in the preparation of manuscripts that are about them. Each student then helped write different sections of the manuscript, particularly with respect to its findings and conclusions. Throughout this process, I took the lead in writing, editing, scheduling and facilitating meetings, and mentoring the student co-authors. Dr. Jezierski provided invaluable support and guidance. Our plan is to submit the chapter to the *Journal of the Scholarship of Teaching and Learning*.

Chapter 5 concludes the dissertation with a summary of implications and contributions of the chapters to the field of data literacy and data science education, mathematics education, interdisciplinary studies, teacher education, and the scholarship of teaching and learning. I also discuss my learning as an educator and researcher and share my plans for further teaching and research.

REFERENCES

Augustine, S. M. (2014). Living in a post-coding world: Analysis as assemblage. *Qualitative Inquiry, 20*(6), 747-753.

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D.A. (2020). *PreK-12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education*. American Statistical Association and National Council of Teachers of Mathematics.

Barton, B. (1996). Making sense of ethnomathematics: Ethnomathematics is making sense. *Educational Studies in Mathematics*, 31, 201-233.

Boylan, M. (2017). Towards a mathematics education for ecological selves: Pedagogies for relational knowing and being. *Philosophy of Mathematics Education Journal, 32*, 1-23.

Bozalek, V., & Zembylas, M. (2017). Diffraction or reflection? Sketching the contours of two methodologies in educational research. *International Journal of Qualitative Studies in Education, 30*(2), 111-127.

Chen, G. A. (2023). Teacherly response-ability: ethical relationality as protest against mathematical violence. *Educational Studies in Mathematics*, 114, 277-296.

Chun, W. H. K. (2021). *Discriminating data: Correlation, neighborhoods, and the new politics of recognition*. The MIT Press.

DataScience4Everyone (n. d.). About the coalition. *DataScience4Everyone*. Accessed at: https://www.datascience4everyone.org/about

D'Ambrosio, U. (1985). Ethnomathematics and its place in the history and pedagogy of mathematics. *For the Learning of Mathematics, 5*(1), 44-48.

de Freitas, E., & Sinclair, N. (2014). *Mathematics and the body: Material entanglements in the classroom*. Cambridge University Press.

Dominguez, H., Abreu, S., & Peralta, L. M. (2023). Young philosophers: Fifth-grade students animating the concept of space. *ZDM*, 55, 1151-1171.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Ingold, T. (2011). *Being alive: Essays on movement, knowledge and description*. Routledge.

Klein, J. T. (2010). A taxonomy of interdisciplinarity. In R. Frodeman, J. T. Klein, & C. Mitcham (Eds.), *The Oxford handbook of interdisciplinarity* (pp. 15-30). Oxford University Press.

LaMar, T., & Boaler, B. (2021). The importance and emergence of K-12 data science. *Phi Delta Kappan, 103*(1), 49-53.

Lockhart, P. (2009). *A mathematician's lament: How school cheats us out of our most fascinating and imaginative art form*. Bellevue Literary Press.

McQuillan, D. (2018). Data science as machinic neoplatonism. *Philosophy & Technology, 31*(2), 253–272.

Mertala, P. (2020). Data (il)literacy education as a hidden curriculum of the datafication of education. *Journal of Media Literacy Education, 12*(3), 30-42.

National Academies of Sciences, Engineering, and Medicine (NASM) (2018). *Data science for undergraduates: Opportunities and options*. The National Academies Press.

Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Palmer, A. (2011). "How many sums can I do"?: Performative strategies and diffractive thinking as methodological tools for rethinking mathematical subjectivity. *Reconceptualizing Educational Research Methodology (RERM)*, *1*(1), 3-18.

Saldaña, J. (2021). *The coding manual for qualitative researchers*. SAGE publications.

CHAPTER 2: A DIFFRACTIVE READING OF AGENTIAL REALISM, DATA SCIENCE, AND NOISE

Introduction

Data are increasingly shaping and organizing everyday life. Seller and Thompson (2016) describe this in terms of "datafication," whereby systems shape data into information that is then fed back into the processes from which such data emerged. As a result, as Seller and Thompson argue, data play a growing role in how we understand ourselves and each other and the actions we take as a result. At the same time, because of the increasing role that data play in politics, in the workplace, and throughout our relationships with ourselves and others, there has been a strong interest in recent years on expanding educational offerings in data literacies[2] and data science (collectively, "data science education") across formal and informal learning settings (e.g., Bargagliotti et al., 2020). These converging trends suggest the importance of attending to the goals and purposes of data science education and how these goals and purposes are conveyed to students. This paper addresses the former concern through the following questions: What norms and values are being suggested about data within data science education reform efforts? What might be some alternatives?

In this conceptual paper, I address these questions by drawing on theories and studies within and outside the literature on data science education. In the first part of the paper, I draw on Van Wart and colleagues' (2020) idea of scripts and valences to examine the epistemological and ontological assumptions about data science and data science education found within

[2] Following the work of scholars such as Bhargava et al. (2015) and Rubel et al. (2022), this paper will use the term data *literacies* because the idea of a singular notion of *literacy* too often promotes an "overly mechanistic view of the world and its problems" (Bhargava et al., 2015, p. 8) and because "[h]istory sheds light on how defining and promoting literacy has been often entrenched with the constructs and perpetuation of power structures within society" (Bhargava et al., 2015, p. 5).

prominent data science education reform efforts. I then spend the remainder of the paper framing these scripts and valences using the ideas of signal and noise (Crawley, 2017; Serres, 1982; Thompson, 2017), which I argue are generative metaphors with which to understand data science and data science education. I use the idea of noise, in particular, to provoke new questions for data science and data science education based around the liminality of the more-than-human world, the interruptive nature of data science and data science education, and the intra-active dynamics between signal and noise that play a significant part in how we think about data science and communicate ideas about data within data-intensive learning environments.

This paper uses the word "data" to mean numbers situated within particular contexts, along with "dynamic, complex, highly structured (or unstructured) collections of pictures or sounds" (Bargagliotti et al., 2020, p. 6). Despite what appears to be a relatively straightforward definition of data, the term data is "a fluid concept with no universally accepted meaning" (Mertala, 2020, p. 33). Etymologically, the term is plural for the Latin word "datum," which means "given." Many people continue to see data as objective, politically neutral, and capable of offering people the ability to understand a phenomena from a detached, distant, disinterested, and disembodied viewpoint (D'Ignazio & Klein, 2020). In contrast, I understand data not to be apolitical or inherently authoritative. Following scholars such as Dixon-Román (2017), McQuillan (2018), and Walter and Andersen (2013), I take the position that in order to understand data more fully, we must also consider broad questions about what it means to be human in a society of increasing datafication, how data implicate the more-than-human world, and how racialized, gendered, classed, and abled bodies materialize and are materialized through data. This paper not only surfaces and provides an alternative for dominant scripts and valences

of data science and data science education, but also considers how scripts and valences for data science speak to these broad questions about living and being in the world.

Scripts and Valences in Data Science Education

Recent developments in the areas of data science and data science education have not occurred in a vacuum. Educational reform efforts around data science shape and are shaped by existing narratives about data science and other fields such as computation, mathematics, and statistics. The analytical lens of scripts and valences (Van Wart et al., 2020) provides a useful way of understanding these narratives. In a study that engaged young people in participatory data projects, Van Wart and colleagues (2020) synthesized the notion of scripts, counterscripts, third spaces (Gutierrez et al., 1995), and social valences (Fiore-Gartland & Neff, 2015) to make sense of the stories that students were telling about data and how these stories emerged. Scripts are "familiar and comfortable ways of knowing and acting in the world, which vary across contexts, cultural communities, and power relations" (Van Wart et al., 2020, p. 129). Social valences of data are the "common expectations and values that explain how and why people gather, interpret, and marshal data toward particular goals" (Van Wart et al., 2020, p. 129). Van Wart and colleagues examined several scripts and valences about data that emerged throughout their study's participatory projects, which they organized along three themes: discovery (data can lead to new insights), actionability (data can be used to improve our lives), and truthiness (data are authoritative). These "scripts about data tend to be optimistic and ahistorical, and typically do not attend to the broader social, political and historical contexts that also underlie action and change" (p. 131). Van Wart and colleagues did not claim to have created a comprehensive list of scripts and valences about data. Instead, they based their analysis on the relevant literature and their

study involving high school students engaging with real-world mapping and community-based data and presenting their findings to authority figures.

In this paper, I direct these notions of scripts and valences toward the *PreK-12 Guidelines for Assessment and Instruction in Statistics Education II* (GAISE II, Bargagliotti et al., 2020), which is a prominent U.S. policy document that is indicative of one approach to data science education reform. I chose GAISE II because it often arises within conversations about data science education reform, particularly within U.S. contexts, and because of its endorsement by prominent organizations including the National Council of Teachers of Mathematics (NCTM) and the American Statistical Association. GAISE II sets forth a framework of essential statistical concepts and provides 22 examples of sample lessons aimed at supporting students as they "gain an appreciation for the vital role of statistical reasoning and data science, and to acquire the essential life skill of data literacy" (Franklin & Bargagliotti, 2020). Data science education reform efforts such as GAISE II promote visions of a desirable data-driven future that support and are supported by advances in big data and computational technologies. The more advances are made in big data and computation, the greater that these reform efforts seek to elevate the importance of data science and data science education. Conversely, reform efforts may have a downstream effect, insofar as changing scripts and valences about data at the K-12 level may eventually shape scripts and valences about data at the higher educational or professional levels as students progress in their educational and professional careers. The authors of GAISE II are not alone in advancing DSE in elementary and secondary education. For instance, the University of Chicago Center for Radical Innovation for Social Change has helped organize Data Science For Everyone, an ongoing national movement to advance data literacies in K-12 education. Their website, the header of which reads "The Data Revolution is Here. Now We Just Need to Teach

It," includes curricular resources and advocacy packs for making data science education a more central part of K-12 curricula. GAISE II and Data Science For Everyone, in turn, are part of a broader ecosystem of reform efforts and initiatives to increase the presence of data science in schools both at the K-12 level and within higher education.

I argue that the scripts and valences of data promoted by GAISE II and Data Science For Everyone tend to frame the nature and purpose of data in terms of at least three future-oriented concerns: workforce preparation, risk management, and surface-level calls for access and opportunity. I briefly discuss each concern in order to elaborate on the scripts and valences around which these and other data science education reforms are developing and for which this paper offers an alternative narrative. Importantly, I do not claim to have engaged in an exhaustive or comprehensive review of the scripts and valences found within GAISE II and Data Science for Everyone. The scripts and valences that I briefly surface and discuss are merely a selection of topics for critique. It is my view that GAISE II and Data Science for Everyone have done important work in elevating the need for data science throughout K-12 curricula, and there is much to recuperate and value within them apart from the scripts and valences that I discuss. I direct the critiques in this section toward the danger of overemphasizing workforce preparation, risk management, and surface-level calls for data science for all students, such that these narratives become the dominant scripts around which data science education reform efforts gather.

Data science education reform efforts tend to promote visions of a data-savvy workforce of the future. The language of workforce preparation appears multiple times within GAISE II. In the preface of GAISE II, its authors claim, "Today, many sectors of the economy and most jobs rely on data skills" (Bargagliotti et al., 2020, p. 1). Citing statements by the prominent economist

Steve Levitt and reports by the Business Higher Education Forum, GAISE II elevates the importance of data-related careers in today's society and creates the impression that such careers will remain valuable in the future. The implied message is that because skills with data are essential in today's labor market, and because GAISE II provides newly updated resources for students to learn these skills with data, then GAISE II is essential toward ensuring that students remain competitive in the workplace. This is suggested to be true not only within the United States but also in a global economy where, in the words of U.S. Senate Majority Leader Chuck Schumer, "whoever wins the race to the technologies of the future is going to be the global economic leader with profound consequences for foreign policy and national security as well" (Flatley, 2021, para. 3). By elevating workforce preparation discourse, GAISE II advances a goal of using data to secure access to participate in dominant systems of power, particularly by treating data as a tool to make decisions and find meaning within mainstream social and institutional contexts (Rubel et al., 2021).

Data science reform efforts reproduce a vision of the future where people use data-related skills to understand and manage the risks of climate change, economic collapse, and other social issues. GAISE II, for instance, suggests the importance of using data-driven skills to identify, anticipate, and manage "pressing world issues" (Bargagliotti et al., 2020, p. 1). GAISE II is part of a network of policy documents, initiatives, news and research articles, and other efforts aimed at using data science education to prepare students and society at large for the risks of an uncertain world. In a practitioner journal published by NCTM, LaMar and Boaler (2021) warn readers that "global crises presented by the coronavirus, the nationwide movements to address racism and racial injustice, the climate crisis, and the tumultuous 2020 presidential election all served as painful and pressing examples that we, as a society, urgently need to equip every

citizen with the skills of data literacy" (p. 52). Social risks are also associated with data practices themselves. GAISE II, along with many other policy and curricular documents about data science, warn against "misleading graphical representations and limitations of data sets" (Bargagliotti et al., 2020, p. 7). It is the ethical responsibility of students to become data savvy enough to avoid creating such "misleading graphical representations" for others and to avoid becoming misinformed by data sets that are misleading or limited.

The approach to risk mitigation promoted by data science reform efforts emphasize individual responses to risk. The "risk society" is a theory that has been proposed by the sociologist Beck (1992) to describe how modern society is organized around socially manufactured risks and their management, as in the case of climate change or economic collapse. Beck argues that risks such as these do not exist on their own but instead emerge from the very acts of measurement and representation that we use to define and communicate risk. Data science reform efforts not only tend to ignore how data practices co-construct the risks they purport to measure from afar, they also place the burden of risk mitigation onto individuals. Central to Beck's concept of the risk society is the concept of "individualization." In a risk society, individuals are expected to use scientific and social scientific principles to identify, manage, and guard against personal risks (Beck 1992; Van Loon, 2002). For instance, a person may be expected to purchase home insurance for themselves in an area with increasing climate change risks, in lieu of more urgent collective responses to address climate change. As I discuss in greater detail in a subsequent paper, data science education reform efforts risk reproducing this individualization by shifting the burden of responsibility for managing risk onto newly data literate individuals. GAISE II states, "Good data sense is needed to easily read the news and to participate in society as a well-informed member" (Bargagliotti et al., 2020, p. 1). Such "good

data sense" is a matter of self-determination, and a critical aspect of self-determination is positioning oneself in the world in order to best anticipate and respond to the uncertainty of the future. In other words, a prominent script promoted by data science education reform is for students to use data science to mitigate the harmful effects of an uncertain world because, at the end of the day, they are on their own.

Prominent data science education reform efforts promote a vision of the future where all students have access and opportunities to learn about and participate in traditional data practices. The authors of GAISE II advocate for "statistical literacy for all" and base their argument for access and opportunity on the increased importance that data play in our world. The Data Science for Everyone initiative seeks to advance DSE "so that every K-12 student is equipped with the data literacy skills needed to succeed in our modern world" (DataScience4Everyone, n. d.). The Data Science for Everyone website continues, "Equitable access to data science education is an opportunity to open doors to higher education, high-paying careers, and an engaged community." In 2020, a group of more than 160 individuals and organizations organized under Data Science for Everyone co-signed a letter to the Biden administration to "encourage states and school districts to incorporate data science into math standards and curriculum" and engage in various "data science for all" initiatives. One of these initiatives included collecting information on data science offerings because "while studies show that there is inequitable access to rigorous math and science courses in high school, there is little information about access to courses that integrate or focus on data literacy." The Data Science for Everyone website also includes a letter sent by ten U.S. senators to the National Science Foundation and Institute of Educational Sciences to "increase access to high-quality data science education, so that all students can keep pace with our 21st century world." The stated goal of

data literacies and data science for all students resembles the incrementalist orientations of the white imaginary as critiqued by Martin (2019) in the context of mathematics education reform. In the white imaginary of data science for all students, similar to the white imaginary of mathematics for all, everyone has access to data-related skills regardless of who they are. The purpose of data science education reform is to give everyone access to and opportunities to participate in existing data practices, just as Martin (2019) argued that discourses about mathematics for all have involved "assimilation into the existing cultures of mathematics education, thereby sustaining the fundamental character of the domain" (pp. 460-461). Reform efforts do not encourage students to question the "21st century world" that such efforts seek to prepare them for. GAISE II, among other reform efforts, is "framed as individual self-protection" (Rubel et al., 2021, p. 219) that reifies rather than transforms systems of power.

Data science education reform efforts may be a well-intentioned attempt to update traditional mathematics and statistics curricula and ensure that students are prepared to participate in the labor market and use data to anticipate and manage problems in the future. Moreover, many of these efforts acknowledge the importance of ensuring that data science education opportunities are equitably distributed. However, there is a tendency within calls to expand data science educational offerings to uncritically endorse data science without sufficient attention to questions of ethics and power. The importance placed on data-related careers is neither neutral nor inevitable (D'Ignazio & Klein, 2020), thereby suggesting that the decision to elevate data science within educational spaces is neither neutral nor inevitable. The script that elevates data science over other careers is supported by a network of metaphors of data scientists as "unicorns," "wizards," "ninjas," and "rock stars" who are capable of using data in rare, magical, and complex ways in order to solve problems (D'Ignazio & Klein, 2020). D'Ignazio and

Klein (2020) argue that these scripts about data scientists are misguided and ignore the data pipeline of unacknowledged and underpaid workers who are also responsible for many of the developments made in data science. Moreover, while a common script in data science education is to prepare students for a risk-filled future, missing from these visions are collective and relational responses to risk and uncertainty. Further missing from these visions is an emphasis on the fact that data emerge from the efforts of people and materials, during a particular time, in a particular place, and for particular purposes, an idea that Loukissas (2019) describes as all data being local. Collective and relational responses to risk and uncertainty, alongside the local nature of data, raise questions about what should be the role of data in how we think about the future and how could we reconceptualize data in terms of our relational rather than individual activity. Finally, surface-level calls of access and opportunity may not be enough to combat the ways that data-driven systems continue to harm traditionally marginalized people on a systemic level. Providing everyone with data-related skills, without an accompanying emphasis on power and ethics, may and likely will merely reproduce the status quo.

These scripts about data science and data science education resemble long-held discourses found within debates about STEM (see e.g., Darville & Disare, 2016). Data science education reform efforts, however, differ from other STEM debates in how they tend to rely on and reproduce a particular "instrumentalization of futurity" (Bahng, 2018, p. 7). In the remainder of this paper, I argue that this instrumentalization of futurity can be framed in terms of tacit desires to elevate signal over noise. Before doing so, I briefly discuss literature that already adopts critical perspectives with respect to data science education. Indeed, data science education reform is broad and encompasses a variety of approaches that do not necessarily conform to the traditional scripts and valences about data that I have discussed thus far. These approaches make

a valuable contribution to the data science education literature by offering a range of narratives about data science education that are an alternative to those promoted by dominant data science education reform efforts such as GAISE II. This paper does not engage in a systematic review of these innovative approaches in data science education (for an example of such a review, see e.g., Lee et al., 2022). Instead, I seek to build on this critical data science education literature by taking a theoretical approach that is less common within such body of work. Although it is not the focus of this paper to dwell on the specifics of policy documents in data science education, highlighting the work that has already begun in this area is an important step toward situating my analysis within existing conversations in the field.

Critical Perspectives About and Relational Ontologies of Data, Data Science, and DSE

Recent years have seen an increase in literature adopting a critical stance on data science education. This literature surfaces and challenges assumptions about data science and data science education, foregrounds the sociopolitical dimensions of data science, and provides examples of such critical approaches within formal and informal learning contexts. Van Wart and colleagues (2020), for instance, ground their discussion of data science scripts and valences in a participatory mapping project where high school students measured, in one case, air quality sensor data and presented their findings to relevant stakeholders. The study pointed to the generative possibilities for student learning around what Fotopoulou (2021) and Stornaiuolo (2020) characterize as critical data literacies, which can be used to nurture student agency with data and sensitize students to issues of power and ethics within data science (D'Ignazio & Klein, 2020). The idea of critical data literacies resonates with Lee and colleagues' (2021) notion of taking a humanistic stance with respect to data science, which includes attending to the sociotechnical dimensions of data science. This goal of attending to the sociotechnical

dimensions of data science is amplified by Hardy and colleagues (2020), who advocate for an approach to data science education that embraces students' engagements with the materiality of the tools used throughout data-driven processes, which they describe in terms of "material resistances" (p. 109). More broadly, data do not enable people to understand the world from an inherently objective and disinterested vantage point, contrary to the belief held by some people (D'Ignazio & Klein, 2020). Instead, as Walter and Andersen (2013) argue in their book on Indigenous statistics, data have a performative effect on the world, particularly when it comes to how we regard and act with respect to nation-states. They write:

> Population statistics in particular are an evidentiary base that reflects *and* constructs particular visions considered important in and to the modern state. They map the very contours of the social world itself. They shape and thus create the accepted reality of things most of us think they merely describe. (p. 7)

In short, the diversity of novel and critical approaches to data science challenges the idea that there is a single "data literacy" made up of a pre-established list of primarily technical, individually held skills. Indeed, such singular notions of data literacy have been used to reinforce inequities and assumptions about who is and who is not data literate (Bhargava et al., 2015) in addition to missing perspectives that connect data to power, ethics, material resistances, nation-states, and other sociomaterial and sociopolitical issues that are inextricable elements of the phenomenon of data.

These perspectives also highlight the reality that contemporary data practices are becoming increasingly entangled with complex material, social, and political relations. This reality makes data science and data science education well-suited to be read through the ideas of agential realism and noise. Scholars such as Dixon-Román (2017), McQuillan (2018), and Sanches and colleagues (2022) have taken important steps to propose and discuss an agential realist account of data science. Dixon-Román (2017) argues that data are haunted by

sociopolitical and material relations of racialization. Drawing on the work of Wynter (2003) as well as theories of hauntings by Barad (2010), Derrida (1994), and Gordon (1997), Dixon-Román considers the specters and ghosts that are beyond what we inquire about through the collection and analysis of data. Data are assemblages "formed, shaped, and entwined with the histor(icit)y of the human" (p. 50). Data produces and is produced by the biopolitical narratives of humankind's ontological origins and the processes by which bodies are materialized as racialized. McQuillan (2018) argues that modern data science is an "echo of neo-platonism" (p. 2) that "creates the structural conditions not only for specific injustices caused by bad data or false positives but also the elevation of epistemic injustice, where data science has more sway than the testimony of the subject" (p. 10). Epistemic injustices become amplified by data science's use and reliance on computational technology, whereby "algorithmic predictions become forceful at a human level" (p. 11). McQuillan draws on agential realism to provide a "non-dualism that contrasts starkly with the current neoplatonism and the possibility for participatory agency" (p. 16).

People can use these ideas to push a relational understanding of data and data science beyond seeing the non-human world as mere mediational contributors to a primarily anthropocentric cognitive activity. The agentic force of matter, which is an ever-present part of its materialization, reinforces the fact that engaging with data is a matter of "the world's worlding itself" (Barad, 2011, p. 133). Response-ability and ethics in data science extends beyond calls for ethics that emphasize accurate interpretation, accurate communication, and the application of data in non-harmful ways. Although the application of data in non-harmful ways is desirable, response-ability is a deeper notion that refers to the capacity to respond, elicit

response, and attend to how our theoretical and action-oriented commitments perform boundary-making practices for which we are always accountable (Barad, 2007).

Building on this aforementioned work, I read data science and data science education through agential realism and, in addition, through the idea of noise. It may appear unusual, at first, to evoke the idea of noise in a discussion about data science and data science education. However, as I will argue, noise has been conceptualized by philosophers in ways that render it a generative metaphor and lens through which to understand data science and data science education in relational terms that complement and build on existing critical and agential realist accounts of data science. I begin by focusing on conceptualizations of noise from two philosophers: Thompson (2017) and Serres (1982). Their formulation of noise as a perturbing force relation and parasite offers an opportunity to surface, confront, and provide alternatives for widespread assumptions about data that remain dominant today. Throughout this discussion, I also draw on Crawley's (2017) philosophical discussion of noise rooted in what he calls Blackpentacostalism, which I argue provides convincing arguments against and alternatives for the normative scripts and valences about knowing and being in the world that are shared across philosophy, theology, and data science.

Reading Agential Realism Through Noise

This section offers a start to a diffractive reading (Barad, 2007) of agential realism through the concept of noise. These ideas, as well as the methodology of diffraction, is grounded in a particular notion of relations. Relations do not exist among people who possess an independent existence and who, on occasion, come into contact with one another. Rather, relations are a matter of humans and the non-human world being co-constitutive of one another (Barad, 2007). Barad uses the term intra-action, as opposed to interaction, to underscore this

point. Diffraction attunes to the idea of intra-action by attending to "relations of difference and how they matter" (Barad, 2007, p. 1). For purposes of this paper, I use the methodology of diffraction specifically to read texts through one another, not in order to compare and contrast texts or to evaluate the strength of one text against the weakness of the other. Instead, following Barad's (2007) own approach to diffractive reading of texts from quantum physics and philosophy, my methodological approach is to bring disparate texts in conversation with one another and attune to what new insights emerge. This approach to reading is premised on the idea that texts, on their own, possess no standalone or inherent meaning. Properties emerge and become coherent on account of diffractive apparatuses of measurement (Barad, 2007); in this case, the diffractive measuring apparatuses are the texts themselves, and what emerges are the "relations of difference and how they matter." Moreover, in the spirit of attending to "fine details" as a "crucial element" of a diffractive methodology (Barad, 2007, p. 92), I point out that the phrase "relations of difference and how they matter" differs from the phrase "differences and how they matter" insofar as the former includes the phrase "relations of." This distinction suggests that relations are not only an essential component of differences, they are on an ontological level what constitutes the very nature of difference. Likewise, relations without difference are arguably not generative of the specific, dynamic, and ongoing separations that give phenomena meaning and shape. In short, a diffractive approach to reading is a matter of attuning to differences in a relational sense.

Barad's (2007) philosophy-physics of agential realism emerges from a diffractive reading of quantum physics experiments and the ideas of scholars such as Neils Bohr, Michel Foucault, and Judith Butler. Agential realism seeks to understand the world not as a collection of objects with pre-existing boundaries and properties but rather in terms of phenomena and the specific,

dynamic, and ongoing separations within them. Perceptions and experiences in the world are the result of boundary cuts made possible by the specificity of discursive-material configurations that render them determinate, give them meaning, and demarcate the lines separating those perceptions and experiences from the agencies who are said to perceive and experience them. Agential realism is not the claim that there is no reality outside of discourse, nor is it the claim that there is some settled reality "out there." Instead, it is a claim of realism that characterizes all matter—humans and more-than-humans alike—as co-constitutive, in constant flux, and mutually entangled. Mutual entanglement is the result of configurations, or apparatuses, that "*are the material conditions of possibility and impossibility of mattering;* they enact what matters and what is excluded from mattering" (Barad, 2007, p. 148). This characterization of apparatuses is suggestive of Barad's overall view of materiality: matter is not so much a noun as it is a verb ("mattering"), implying that matter, materiality, and the physical world move across moments of potentiality and between moments of determinacy and indeterminacy.

Just as Barad's (2007) text on agential realism involved "drawing out the ontological dimensions of Bohr's framework" (p. 174), this analysis seeks to use a diffractive reading of Barad's agential realism with Thompson (2017) and Serres' (1982) notions of noise in order to attune to the ontological possibilities emergent within the data science education literature. For Serres (1982), noise is associated with the term "parasite." This term focuses on aspects of relations that perturb the "regular" function of a system. Serres begins his text on the parasite with an interpretation of a fable involving a country rat feeding at the home of a city rat, who in turn is feeding off a tax farmer. The rats' meal is then interrupted by a noise at the door. "Parasite" translates from its original meaning in French to one of three interrelated meanings in English: biological parasites, social parasites, and informational parasites. Parasites, which can

serve as obstacles to communication and working systems, are also necessary for communication to occur and for systems to function. In Serres's fable, the rats, the farmer, and the noise at the door are all parasites in some way—as an organism feeding on a host (biological), as an uninvited guest (social), and as an interruption in communication or the normal course of business (static/noise). Noise is more than the cacophony that one often associates with the term. Serres's philosophical rendering of noise elevates the generative possibilities of disorder, dysfunction, and perturbation. For Serres, sameness is tantamount to nothingness. Referring to "rationalists" who reject the value of disorder, Serres (1982) states, "He is horrified by the complex. He does not understand that chance, risk, anxiety, and even disorder can consolidate a system. He trusts only simple, rough, causal relations; he believes that disorder always destroys order. He is a rationalist, the kind we just spoke of" (p. 14). In other words, some degree of dysfunction, whether in the context of an individual organism or a larger institution, is always necessary for a system to function, whether such system is a classroom, a school, or an entire field such as data science or data science education.

Thompson (2017) offers a similar conceptualization of noise. She draws on examples from music studies, sound art, and noise performance to define noise as "a perturbing force-relation that, for better or worse, induces a change" (p. 42). Although Thompson draws on examples of noise throughout these specific contexts, she conceptualizes noise in affective terms and emphasizes that noise is not limited to any particular type of sound or source. Instead, noise, as a perturbing force-relation, spans all forms of relations. She states, "[N]oise also acts within and in relation to that which is designated the 'non-human': both noise and affect, and noise as affect traverse the distinctions drawn between organic and machinic, natural and 'unnatural', acoustic and electric, analogue and digital (p. 54). Noise does not need to be perceptible to

human ears to be considered as noise. As Thompson (2017) argues, noise is the vibrational medium from which "signal" emerges within the human and non-human world, thereby challenging distinctions often made between the two. Moreover, noise is not always a positive force for good. Thompson (2017) provides several examples where noise is harmful, such as the weaponization of noise to characterize Black and Brown people as "foreign Others"—as in the case, for instance, of labeling Black and Brown people as "loud" and "unruly" (Coyne, 2020)—or the literal use of noise as an acoustic weapon. The central idea for Thompson is the capacity for noise to affect: "[N]either the noise of the medium nor the noise of sonic weapons can be fully grasped through a consideration of the personal affections of a listening body-as-subject. Rather, they show noise to be affective in the broadest sense – of one entity acting upon another" (p. 51). For Thompson (2017) and Serres (1982), the purpose of attending to noise is to point out and disrupt, among other purposes, normative ideas about the "pure and perfectly transmitted sound signal" (Thompson, 2017, p. 98). In contrast to assumptions about noise as an undesirable phenomena to eliminate, minimize, or control, noise is a necessary third term within any relation. Exactly which element within a relation is operating in the role of "the third" is dynamic, shifting, and sometimes unclear: "Who is the host and who is the guest? Where is the gift and where is the debt? Who is hospitable, who is hostile, again the same word, the same thing" (Serres, 1982, p. 23). In this expansive notion of noise, noise is a paradoxical concept marked by and productive of complexity and ambiguity.

By reading agential realism through noise, it becomes possible to read notions of disruption, interruption (later, "intra-ruption"), static, and parasitic relations through the agential realist notions of diffraction, intra-action, and agential cuts. Under an agential realist account of natureculture, distinct agencies and moments of separability emerge from cuts with/in

phenomena (Barad, 2007). When I first read Barad's work, my initial instinct was to imagine two agentivities that intra-acted with one another, the result from which co-constituted differential boundaries emerged. This interpretation resonated with the examples from quantum mechanics that Barad (2007) provides to illustrate and shape their theory of agential realism. In these examples, an apparatus of measurement exists in a co-constitutive relationship with an object of inquiry. Together, they constitute a phenomena from which the distinction between subject and object becomes momentarily legible. Noise introduces the idea of a "third" within the "ontological inseparability-entanglement of intra-acting 'agencies'" (p. 139). The idea of a "third" infuses the three-part notion of parasite into an understanding of the agential realist notion of phenomena. Certain apparatuses of bodily production work to "interfere" with the formation of agential cuts that other apparata seek to produce. Trajectories become displaced. "Lines of flight" (Deleuze & Guatarri, 1987) become bent. Resonance becomes discord. Thus, according to a noisy interpretation of agential realism, every encounter—for instance, that between a person and data—entails a messiness and dysfunction that corrupts any concept of purity or harmony and leaves in its wake something generative and new.

Put another way, the infusion of noise within agential realism suggests that every intra-action possesses an element of *intra-ruption*. Just as intra-action rejects the idea that relationships consist of two or more independent forces that come into contact with one another, intra-ruption rejects the idea that ruptures occur when an outside force intervenes on an otherwise closed and coherent system. Intra-ruption takes Thompson's idea of noise as a "perturbing force-relation" and Serres's idea of noise as a parasitic third and suggests that the distinction of noise, as such, emerges from and contributes to the enactment of an agential cut. The ambiguity between noise/signal and host/parasite, which Thompson (2017) and Serres

(1982) describe, is well aligned with this idea. Parasites are endemic to functioning systems, and the role of the parasite seemingly shifts from one participant in a relation to another precisely because parasites/static/noise are "part of the ongoing dynamism of becoming" (Barad, 2007, p. 142). In other words, while Barad's use of the term agential "cut" suggests a precise—albeit dynamic, open-ended, and iterative—action through which the inherently indeterminate becomes determinate, the idea of rupture suggests a messiness and a fraying. The determinant is always at the edge of becoming indeterminate (or becoming determinate in a new way) on account of intra-ruptive forces co-constituted within a phenomena. Noise is neither guaranteed to carve a path toward liberation nor guaranteed to reproduce structures of oppression. Instead, through its paradoxical and ambiguous nature, noise possesses the potential for generative change, "for better or worse" (Thompson, 2017, p. 42). This potential for generative change, in turn, entails an accountability to the forms of disruption that it can create, just as Barad (2007) emphasizes an accountability to agential cuts. Noise is a call for attuning to discord and a call to calling into question what exactly one means by "attunement" in the first place.

What does a diffractive reading of agential realism through noise mean for data science and data science education? What might be the implications for teaching and learning within data-intensive learning environments? I argue that a diffractive reading of agential realism through noise through data science is one way to characterize dominant scripts and valences about data science and data science education in terms of widespread desires to elevate "signal" over "noise." Thompson's (2017) and Serres's (1982) philosophical inquiry of noise arises, in part, as a response to the treatment of noise within information theory, a field developed by Shannon and Weaver (1964) in the first half of the 20th century. Information theory concerns the transmission of information from one location to another, whether this information is encoded in

speech, music, images, text, or any other representation. The transmission of information, however, always comes with unwanted additions, such as distortions of sound, static, distortions of shape or picture, or errors in transmission (Shannon & Weaver, 1964). These distortions are characterized as noise. In this view, noises are unwanted, external elements that enter into the path between the sender and receiver of a message, much like debris might enter a pipe and disrupt the flow of water through the pipe. Noise is considered an undesirable characteristic to be juxtaposed against the more desirable concept of "information." Formally, information is the probability of a signal remaining intact as it travels from one location to another. Noise is information's probabilistic foil—it is a randomly generated phenomena that creates undesirable uncertainty in the transmission of a signal. The minimization and regulation of noise is a central idea in information theory and forms the basis of widely used technologies such as barcodes, computer memory storage, and data packet transmission across computer networks.

Noise plays a similar role in dominant quantitative and statistical thought as it does in information theory. Noise is linked to variation and error as these concepts appear in the collection, analysis, and representation of data. It is the unexplained, spurious, and ultimately unwanted variation that all data exhibit. Malaspina (2018), another philosopher of noise, states, "[I]n modern statistics precision itself has become a question intimately linked to noise: Precision is a measure of random noise" (p. 1, citing Smith, 2002). Probability and statistics are tools designed to identify, isolate, and regulate noise so as to discern the "true information" (also called the "signal") found within datasets. This idea of separating signal from noise within data became prominent in the 19th century, where the founders of modern statistical thought sought to discern the "true" characteristics of populations in order to make claims about the superiority of some people (white able-bodied men) over others (everyone else). Malaspina (2018) explains:

> The singularities we now associate with the idea of noise as 'unexplained variation' fail to constitute an 'event' in the eye of the nineteenth-century statistician, as they are subsumed and diffused into the necessary unfolding of the statistical tide of large numbers. So too the singularity of free will, blends itself into the propensities of the masses. What emerges in the 'man without qualities', whose singularity is washed away by the flow of great numbers, subjecting the individuality of human agency to the same laws of statistical physics. (p. 139)

The isolation, regulation, and elimination of noise is part of a broader statistical vision where people and their activities can be reduced to data. In turn, noise within data (i.e., random error) must be separated from what data "are really saying" in order to generate a clear and coherent account of the world as it truly is.

A diffractive reading of noise through agential realism offers the opportunity to reiterate alternative narratives about noise begun by scholars such as Thompson (2017), Serres (1982), Malaspina (2018), and, as I will later discuss, Crawley (2017), while understanding how these alternative narratives might speak directly to data science and data science education. It is important, however, to first acknowledge critiques that scholars have laid with respect to new materialist scholarship, with which Barad's ideas of diffraction and agential realism have been associated. This is essential because of the central role that Barad's theories play throughout this paper. Scholars such as Todd (2016), for instance, have critiqued new materialism's failure to cite the Indigenous scholars whose ideas about interdependence and attention to the natural world long preceded "new" materialism's turn toward the ontological, biological, and material. Ahmed (2008) argues that new materialist scholarship risks gathering around the taken-for-granted gesture that "non-new material" feminisms are anti-biology and overly preoccupied with social constructionism, to the point that new materialist scholars forget or caricature legacies of feminist work. Similarly, James (2019) critiques new materialism's reliance on notions of vibratory resonance to claim that their methods are better alternatives to traditional philosophical

thought. The problem, as James argues, is that "by using acoustic resonance as a model for philosophical theorizing, they create the same relationships among philosophers and their theories that neoliberalism creates among people" (p. 88). Other scholars have responded or addressed these critiques, either by seeking to clarify the claims being made by new materialist scholars (Davis, 2009; van der Tuin, 2008) or promoting anti-colonial engagements between new materialisms and Indigenous literature on agent ontologies (Rosiek et al., 2020). A common question among these critiques and responses is how new materialist philosophy characterizes or fails to consider traditions of intellectual, philosophical, and spiritual thought that precede its and against which new materialism positions itself as something "new."

The issue arises as to how I reconcile the use of new materialist scholarship throughout this paper with the critiques and debates that surround new materialisms. Indeed, a central tenet of Barad's (2007) agential realism is to consider the cuts, enactments, and foreclosures of the theories we produce and use. It follows that accountability extends to the use of new materialist scholarship itself. By citing scholarship such as Barad (2007), I do not claim to use theories that are a better or newer alternative to traditions of feminist, Indigenous, or other philosophical or spiritual thought. The purpose of this paper is not to distinguish new materialist scholarship from other forms of scholarship. Rather, I use new materialisms—and Barad's ideas of diffraction and agential realism in particular—in my analysis because these theories are capable of speaking directly to data science and data science education insofar as they help draw out notions of signal and noise that I argue are an undertheorized tension within data-intensive learning environments. Theories about noise are especially fruitful in this regard, and new materialism offers a strong starting point for such analyses. My use of new materialisms, in other words, is narrowly tailored to the specific context of this study's focus on data science and data science education and not

intended as an analysis that can be readily transferred or generalized to other contexts. Indeed, doing so would be antithetical to new materialism's emphasis on particularity and difference and would only serve to reinforce the critiques laid against this body of scholarship. Nonetheless, despite my intentions in how I seek to use new materialist scholarship, it is equally true that such scholarship and my use of it is marked by some degree of erasure. This is true of any use of theory—and perhaps especially true with new materialisms—and my hope is that my engagement with the ideas in this paper and in subsequent work reflect a commitment to responsible and respectful scholarly work.

The remainder of this paper elaborates on a reading of agential realism through noise through data science in greater specificity. I argue that such a reading can reenable ways to (1) consider the role of liminal agents in data science and data science education; (2) understand data science and data science education as forms of intra-ruption; and (3) understand data science and data science education as the dynamic intra-action between signal and noise. These understandings not only offer a critique of dominant scripts and valences of data science and data science education but also build on alternative narratives for data science grounded in a relational orientation to knowing and being.

Attuning to Liminality: Materiality and Noise

Due to the close ties that data science shares with mathematics, there is a risk that many of the ontological and epistemological assumptions that people hold about mathematics will carry over into how people think about data science and data science education. Mathematics is often treated as a purely abstract pursuit with no bearing on bodies or the material world, an idea that has been challenged by various scholars in mathematics education (e.g., de Freitas & Sinclair, 2013). Although it is more difficult to deny the presence of the material world within

data science as compared to mathematics, the role and agency of matter in data science can at times be similarly downplayed or forgotten. As Sanches and colleagues (2022) emphasize, however, "[D]ata is inextricably connected to the physical and material arrangements that support it, from the physicality of server infrastructures and fiber optics, to sensing mechanisms, and when it comes to data about humans, their fleshly bodies" (p. 2). Even when the materiality of data is acknowledged, there are some views that treat data as akin to a two sided object, with one side consisting of the material dimension of data and the other side consisting of its abstract—and therefore more pure and ideal—counterpart. McQuillan (2018) expresses this idea in his discussion of data science as a form of "machinic neoplatonism": "For the data scientist, computation plays the role of the intermediary between the imperfect world of data and the pure function that relates the features to the target" (p. 9). Statistical analysis often involves the delineation of a "true" parameter that must be estimated using a sample of "imperfect" data burdened by random error, which some practitioners refer to as noise. Here, the materiality of data is not denied. Rather, the agentic role that matter plays in data is treated as a form of unwanted interference. The purpose of data and statistics is to uncover the underlying structures of randomness and variation that govern the universe. Hacking (1990) calls these structures "law-like regularities" (p. 3) and traces their emergence in Western culture through events such as Quetelet's creation of the statistically average man. Just as Platonic idealism in mathematics has sought to uncover the hidden structures of abstraction, and physics has sought to uncover natural laws of the universe, data science has taken up a similar project by seeking control, mastery, and certainty over the notion of uncertainty itself.

New materialisms (e.g., Sanches et al., 2022) and traditions of Indigenous thought (e.g., Ellenwood & Foxworth, 2014), among other traditions, contribute alternative perspectives for

understanding the role of the material world in data literacies/science. Ellenwood and Foxworth, citizens of the Nez Perce Tribe and Navajo Nation, respectively, discuss the connection between Indigenous material culture—such as clothing, horse gear, and other items—and Indigenous data sovereignty. Their discussion concerns the repurchase and return of a large set of stolen items known as the Wetxuuwi'itin' Collection. They argue that such material culture are noit "vestigates of the past" but rather "the embodiment of relational ties and cultural knowledge between those who made them and the Nimı´ipuu today" (p. 131). Citing the work of Indigenous scholar Stephanie Carroll (2024), they state at length:

> Indigenous Peoples' conceptions of Indigenous data are expansive with fluid boundaries of data, information, wisdom, and knowledges, particularly in juxtaposition to colonial contexts that focus on digital characters, quantities, and symbols as data distinct from other forms. Indigenous data can be digital or emerge as tangible or intangible information, knowledges, languages, cultures, resources, materials, specimens, and objects. Indigenous people have always created, stewarded, and collected data. (p. 139)

Ellenwood and Foxworth (2024), alongside Carroll (2024) and other scholars involved in the field of Indigenous Data Sovereignty (IDS), teach us that data do not need to be relegated to mere "digital characters, quantities, and symbols." Instead, data can be living material culture that embody relational ties and transgenerational knowledge. Rather than being inert, data, in its materiality as "oral stories, stories, winter counts, calendar stick, totem poles, and other instruments...store[] information for the benefit of the entire community" (p. 139). Importantly, issues of IDS bring to light explicit connections between data, sociopolitical issues, ethics, and questions of citizenship and sovereignty. The case of the Wetxuuwi'itin' Collection, among many other similar situations, raise questions about the relationship between Indigenous material culture, philanthropy, and colonial institutions such as museums (Foxworth & Ellenwood, 2024).

In a different context, Sanches and colleagues (2022) use agential realism to understand how biodata are "produced by instruments…entangled with the world" (p. 3). They discuss an example of a person using their index finger to measure their pulse. In this analog configuration, "we enact an agential cut between the parts of our body that are measuring and the others that are being measured, which stabilizes the world in a way that allows for a pulse to emerge" (p. 3). This articulation emphasizes the entanglement of data with matter and challenges attitudes about materiality that treat matter as a mere conduit for human intention. In the case of measuring one's own pulse, the data that is collected is both *by* the body and *about* the body, with the two emerging from the same body, but "[i]n this gesture, our own bodies get split into two" (p. 3). Understood through an agential realist account of the world, the gesture of measuring one's own pulse is the intra-action between the apparatus of measurement and the object of measurement. Sanches and colleagues (2022) gesture toward this idea of intra-action when they state, "Data, in this case a heart rate, can be seen as an inscription of that process of mutual orientation and co-constitution" (p. 3). Understanding data creation as an intra-active phenomenon invites a shift away from the idea that we can use tools to objectively measure an external world for personal and economic purposes. Instead of data being used exclusively to tackle pre-defined problems in standardized ways, data in its materiality is seen as a lively, dynamic, and agentic part of the materialization of new problems and what if statements (Sanches et al., 2022). Sanches and colleagues analyze case studies involving biodata designs that illustrate diffractive engagements with data. These case studies include, for instance, a biosensing device to study a professional singer's breathing. As Sanches and colleagues (2022) state, "This case study illustrates an example of engaging with biosensors diffractively, attending to the differences between bodies, the different ways of defining what breathing is, and consequently of designing with and for

breathing" (p. 6). New questions emerge that go beyond merely measuring breathing from a distant and disinterested vantage point. Instead, by attuning to the mutual co-constitution of bodies, breaths, and measuring devices, the researchers and the professional singer provoke new questions about what exactly entails breathing, how it connects to bodily movement, and what can be measured as a result of these questions. As Barad (2007) emphasizes, the creation and approach of such new problems entails accountability. Agential realism can be read as emphasizing a response-able orientation toward data and an attunement to what else is possible—and what is excluded—through data.

These ideas can be carried over into data science education contexts. Hardy and colleagues (2020), for instance, advocate for learning opportunities that allow students to encounter the "material resistances" (p. 109) of data processes. Providing such opportunities, they argue, can spur student agency by requiring students to account for these resistances and understand "data as contingent on material and technological origins" (p. 109). New materialisms can be used to point out that Hardy and colleagues omit the fact that encounters with material resistances also include an ethically-charged accountability to the agential cuts that these resistances produce. Beyond asking students to attend to the material contingency of data, educators can also ask students to consider and challenge power as it relates to data's material dimensions, and to embrace these material and contingent dimensions of data as an act of political resistance against dominant ways of understanding data as a form of disembodied reason (see D'Ignazio & Klein, 2020). Central to embracing the materiality of data in these ways is attunement to how people are inextricably bound up in the meshworks (Ingold, 2011) of materiality that comprise data and data practices. Entanglement entails not only considering

flows of power within data assemblages but also considering how one can and should respond to such entanglements (Barad, 2007).

What can reading data and agential realism through noise offer beyond what has already been suggested by scholarship that imbues a new materialist lens onto data? First, evoking the concept of noise is one way of reframing traditional ideas about data science, which can be said to use mathematics, statistics, and computer science principles toward the goals of eliminating, reducing, or controlling the "noisy" aspects of data. In this framing, noise takes the form of variation produced through "random error" arising in part from the physical instruments used throughout the data process. Efforts to attend to the materiality of data frame matter in terms of mediation, such as in the following description: "Indeed, data are cultural artifacts created by people, and their dutiful machines, at a time, in a place, and with the instruments at hand for audiences that are conditioned to receive them" (Loukissas, 2019, p. 2). The concept of data itself is framed as a mediating device that shapes our reality: "Data increasingly mediates how we understand the world" (Van Wart et al., 2020, p. 127). Replication is an effort to "de-mediate" or "denoise" the corrupting influence that messy data and physical tools have on the "pure function[s]" (McQuillan, 2018, p. 9) that govern reality.

Instead of noise as something to be eliminated, reduced, controlled, or de-mediated, Thompson (2017) offers an alternative conception of noise that imbues the "noisy" parts of data not only with agency but also with generative possibility. Thompson draws on the interconnected fields of sound, music, and noise studies to argue that noise is the substance from and through which sound emerges and is made possible. Noise is not necessarily the din of a busy city street. Drawing on the work of musical composers such as John Cage, as well as the Japanese noise movement known as onkyô, Thompson states that noise can be the background hum, occasional

interruption, and/or vibrational presence within a room that is otherwise deemed to be "silent." Music never comes from silence because it always begins alongside a persistent presence of noise. She states:

> Emergent sound-signals resonate with the background buzz or form patterns of interference with it: they affect and are affected, animate and are animated, by the vibrational plane. Background noise and sound-signal are co-productive: sounds transform the noisy, vibrational medium/milieu as they emerge and return to it, while the noisy, vibrational medium/milieu helps to shape the sound, contributing to its timbral quality and undertone. (Thompson, 2017, pp. 76-77)

Noise, in other words, is generative—not merely in the sense that noise serves as a raw material out of which people can produce meaningful sounds. Instead, noise is *necessarily* generative insofar as signal cannot emerge without noise; signal and noise are unavoidably co-constitutive. The co-production of signal and noise makes it difficult, if not impossible, to distinguish one from the other. This is not to say that distinctions between noise and signal are never made. Distinctions that are made between noise and signal are the enactment of an agential cut (Barad, 2007). These cuts emerge from the entanglement of signal and noise and, as I have previously noted, these cuts are never stable because there is always an ambiguity of who and what constitutes the parasitic "third" in the web of relations that constitute phenomena.

Noise can be said to build on the work of scholars such as Dixon-Román (2017), who conceptualize data as assemblages of discursive-material relations. The concept of noise implies that these relations are always parasitic. Relations feed off and interfere with one another, and it is never clear who is the guest and who is the host (Serres, 1982). Thinking about data parasitically, for instance, can provide a new way of thinking about the process of generating data (widely referred to as "data collection"). This process often entails the use of instruments of capture such as video and audio recording equipment. Noise invites us to ask how these material agencies interfere with, in some cases, and enable, in other cases, ways of knowing and being.

Likewise, we might ask how our own human presence and actions interfere with, disrupt the trajectory of, and enable instruments of capture. Finally, we might ask how these relations, taken together, form a parasitic "third" in relation to the people, places, and objects from whom the data is "collected." These considerations are a departure from Loukissas's (2019) characterization of data as "cultural artifacts created by people, and their dutiful machines" (p. 2), not only insofar as they challenge humanist assumptions about data-driven practices but also undermine the notion that any machine can ever be "dutiful." Reading noise through data reminds us that our intra-actions with the material world are marked not just by the agentic capacities of the non-human world but also by relations that are asymmetrically power-laden and at times inharmonious.

It is important to reiterate that noise is not necessarily a positive force-relation in data science or data science education. While students can be provided opportunities to attend to the material contingency of data, they can also be provided opportunities to understand that this contingency is a matter of ethics and that ethics is often a matter of ambiguity. For instance, students can be invited to consider when elevating signal over noise is beneficial in some circumstances and harmful in others. D'Ignazio and Klein (2020) argue that data science practices too often elevate the "god trick" (Haraway, 1988), or the idea that data can enable humans to understand phenomena from a detached and disembodied perspective. However, they also implore readers with the following statement: "Don't Never Do a God Trick" (p. 91). They state, "Even though the god trick can do harm…there are also good reasons to use the god trick as a form of recuperation, contestation, or empowerment" (p. 92). In a similar way, Thompson's (2017) characterization of noise as an affective force-relation that induces change "for better or worse" suggests a similar call: Don't Never Elevate Signal Over Noise. At times, there are good

reasons for students to learn data practices that use traditional analytical tools and practices from mathematics, statistics, and computer science. Weiland and Williams (2023) examine how "culturally relevant data" (p. 4) can be used within statistics and data science classrooms to engage students not only in learning more about statistical and data practices but also learn more about themselves and their world. They provide an example, which one of the authors implemented in their own statistics classroom, of using a dynamic data visualization to engage students with the topic of racial segregation in the United States. The students were asked to consider census tract data in New York City and other major cities and apply a Chi-Square Goodness of Fit test to test a hypothesis about racial segregation in those cities. The authors found that the lesson was one of the favorite lessons among many students in the course, with some students stating how data provided new and interesting insight into the issue of racial segregation within cities. In this example, statistics was used to elevate signal over noise as a means to engage students with a topic that was meaningful to them.

Problems can arise, however, when there is an exclusive focus on traditional data analytic techniques as the sole means of knowledge production. This exclusive focus may arise from and be reinforced by the belief that statistical and computationally-based analysis are superior ways to know and understand the world. In the example about racial segregation in New York City, Weiland and Williams (2023) argue that a lesson using culturally relevant data should not have student development of statistical techniques as its sole pedagogical goal. Instead, a culturally relevant approach to statistics and data science education can serve two simultaneous pedagogical goals: the development of statistical acumen and a greater understanding of the context under investigation. They state, "The activity students are engaged in with a CRD should include a high level of interaction with the context, where students must use their data

literacy skills to learn more about the context itself, or make decisions in the context, based on the data" (p. 4, citing Weiland 2017). Weiland and Williams (2023) suggest a tempering of the elevation of signal over noise and encourage educators to provide ample opportunities for students to "read beyond or behind the data" (p. 4). In other words, they suggest the value of striking a balance between elevating signal and attending to noise. Noise, in this case, takes the form of real world context that some statistics or data science educators might argue renders data excessively messy and therefore inappropriate for a learning setting, or serves as a distraction to the work of learning about statistical and data-driven techniques. This type of noise can be embraced as highly relevant to processes of data analysis, and an explicit embrace of this type of noise can be incorporated into task design and student discussion.

Noise can be used to push this example from Weiland and Williams (2023) further. The association that Thompson (2017) and Serres (1982) draw among noise, bodies, and the material world invites the following questions: How do the materialization of bodies—both in terms of the instruments used to generate, analyze, and communicate the racial segregation data, as well as the bodies in the classroom along with the bodies that have been labeled along racial lines by the U.S. Census Bureau—intra-act with one another to co-constitute the data-rich learning environment? How are these forms of noise generating or intra-rupting various ways of knowing and being in the classroom? How might the data-driven conclusions from the class activity be used to take action, such that the conclusions can serve as the "perturbing force-relation" (Thompson, 2017, p. 42) of noise to induce change within one's own life or community? How might the use of data-driven action create ruptures of otherwise possibilities "for better or worse" (Thompson, 2017, p. 42)? Finally, who and what are serving as the parasitic intra-ruptions within the meshwork of relations that constitute the data-driven inquiry? These questions highlight the

generative potential of a deep, philosophical investigation of noise to draw out the liminal agencies that are always present within data-rich learning environments.

Data is not a guaranteed pathway for clear and coherent knowledge production because knowledge is rarely clear and coherent in the first place. Instead, classrooms can serve as a space that challenges common scripts about the importance of elevating signal over noise and invites students to consider noise as an unavoidable, generative, and ethically ambiguous source of intra-ruption. The next section takes up the issue of data science and data science education as a form of intra-ruption with respect to other ways of knowing and being in the world.

Data Science and Data Science Education as Forms of Intra-ruption

In its report for data science education for undergraduate students, the National Academies of Science and Medicine discuss the importance of supporting students in engaging in ethical data practices. These ethical considerations consist of "deciding what data to collect, obtaining permissions to use data, crediting the sources of data properly, validating the data's accuracy, taking steps to minimize bias, safeguarding the privacy of individuals referenced in the data, and using the data correctly and without alteration" (2018, p. 3). NASM (2018) proposes the codification of data science ethics into an oath similar to the Hippocratic Oath in the medical field. The report presents a draft form of such an oath (p. 118), the first several "covenants" of which I reproduce in Figure 1 below:

Figure 1. Data Science Oath

Data Science Oath

I swear to fulfill, to the best of my ability and judgment, this covenant:

I will respect the hard-won scientific gains of those data scientists in whose steps I walk and gladly share such knowledge as is mine with those who follow.

Figure 1 (cont'd)

> I will apply, for the benefit of society, all measures which are required, avoiding misrepresentations of data and analysis results.
>
> I will remember that there is art to data science as well as science and that consistency, candor, and compassion should outweigh the algorithm's precision or the interventionist's influence.

I argue that this oath is emblematic of a preoccupation within data science and data science education to elevate signal over noise. This preoccupation is more nuanced than the belief that data is politically neutral and objective. Instead, data science is framed as a matter of "benefit society" by "avoiding misrepresentations of data and analysis results," a statement that positions data science as a field that can be reigned in, controlled, and directed for benevolent purposes, as long as one approaches their work with "consistency, candor, and compassion" and does not give in to "the algorithm's precision or the interventionists's influence." These statements not only can be interpreted as promoting anthropocentrism through data science but also can be interpreted as aligned with neoliberal fantasies of clarity and control through measurement practices. These human-centered neoliberal fantasies are the elevation of signal, whereas noise can be said to be data's "excess," which cannot be captured, controlled, or regulated through a well-intended commitment to ethics as codified within an oath.

It is worth noting that this paper's discussion of the NASM report or its data ethics oath is not intended as an all-out critique. The oath is a start toward providing nuance to perspectives that continue to characterize data science as a simple matter of applying techniques to numbers without careful attention to context or data's "unintended societal consequences, such as inequality, poverty, and disparities due to algorithmic bias." Framing data ethics in terms of signal and noise is one way of drawing out the power-laden, messy, and parasitic nature of data science that is often ignored in conversations about ethics.

D'Ignazio and Klein (2020) critique data ethics for not going far enough to identify and challenge the underlying problems of data science, which find their roots in what Collins (2000) refers to as the matrix of domination. Data justice includes, among other things, challenges visions of a data-filled future where everyone with data-related skills are adequately compensated, recognized, and valued for their work. In reality, much of the work that goes into creating data-driven products—such as data visualizations, business and consumer facing software, and research studies—remains invisible and unseen in the eyes of the people who use these products (D'Ignazio & Klein, 2020). The data pipeline consists of invisible, underpaid, and undervalued labor, which is also structured along lines of race, gender, and class (D'Ignazio & Klein, 2020). D'Ignazio and Klein (2020) put forward their data feminist principle of making labor visible to extend into data science long-held feminist concerns about the invisibility and lack of value attributed to work that is traditionally coded as female. Just as capitalist societies paradoxically undervalue yet rely on the care work and reproductive labor of women, data-driven societies undervalue yet rely on the invisible labor of people who collect and curate vast amounts of data. These invisible data laborers, particularly in the Global South, tend to be women and/or people of color (D'Ignazio & Klein, 2020).

Jon (2022) similarly discusses the role of women of color in STEM. In addition to noting the experiences of marginalization that many such women experience, Jon characterizes Black feminists in STEM as "subaltern academics" (p. 2) who offer "a unique perspective on systems of power" and "practices beyond the dominant culture" (p. 2), such that "the world can be seen in a radically different way through the window of the most marginalised, as their buried stories *reveal* what our present system of knowledge prevents us from knowing what's possible beyond the mainstream" (p. 9). Jon (2022) explains that such "liminal agents" are "not just victims of

modernity but creative re-appropriators" (p. 14) of scientific and technological practices. This position resonates with Gaskins's (2021) Techno-Vernacular Creativity, which is a framework for culturally responsive STEM practices involving the remixing, reappropriating, and inventive use of technology by people and communities from traditionally marginalized backgrounds.

From this perspective, noise takes on a meaning that goes beyond the liminal agentic capacities of the more-than-human world. Noise can also refer to the people who have been positioned as such. Noise, as Thompson (2017) states, has been associated with "bodies marked as 'other'," along with "the characterization of 'foreigners' as 'noisy', and numerous stereotypes of poor and/or working classes as 'rough', 'brash', 'loud'" (p. 27). Just as traditional approaches to data science seeks to minimize or control the informational and material dimensions of noise in favor of elevating signal, the marginalization of people—particularly women of color—within data science can be interpreted along similar lines (D'Ignazio & Klein, 2020). Jon (2022) and Gaskins (2021), however, highlight how noise is not merely that which mainstream society seeks to subdue but also how noise is a "perturbing force-relation that, for better or worse, induces change" (Thompson, 2017, p. 42). While the underpaid, invisible labor within the data pipeline (D'Ignazio & Klein, 2020) underscores how data science is itself a form of intra-ruption to some people's ways of living and being, the unique perspectives and life experiences of people who have been marked as "foreign Others" can give rise to intra-ruptions and otherwise possibilities for how one can use data for goals of co-liberation and collective empowerment. In the face of data science ethics oaths that suggest the importance of "avoiding misrepresentations of data and analysis results" under the tacit belief that data is a path toward clarity "for "the benefit of society," a reading of data science through noise invites the question that Serres (1982) asks about all relations: "Who then is the real interrupter?" (p. 14). Data science classrooms can be

spaces where students are asked to consider what and *who* are being intra-rupted through data-driven practices. What possibilities are enabled by data, what possibilities are cut off by data, and what are our corresponding response-abilities to such intra-ruptions?

This discussion is not intended to perpetuate stereotypes about people of color as noisy and disruptive. Noise is a broad term that does not merely refer to moments of cacophony. Moreover, cacophony need not be associated with something negative or undesirable. Silence can be oppressive, as in the case of a gentrified neighborhood where norms about quietude are often weaponized against existing residents who have lived in and built up the community across multiple generations (Gonzalez, 2022). Crawley (2017), who James (2019) cites as an example of sound studies scholarship "calibrated to the epistemic, ontological, aesthetic, and political practices black people have used to build alternative realities amid white supremacist patriarchal domination" (p. 6), makes an explicit connection between noise and racialization in his work studying the aesthetic practices found in what he calls Blackpentacostalism. By Blackpentacostalism, Crawley (2017) means "a mutiracial, multiclass, multinational Christian sect that finds one strand of its genesis in 1906 Los Angeles, California" (p. 4). These aesthetic practices include worship practices of church attendees, such as practices of tarrying and testimony. For Crawley, these practices are acts of resistance against the Western theological and philosophical traditions of creating categorical distinctions and hierarchies, which Crawley argues are always racialized. He states:

> [T]he choreosonic performance of Blackpentacostalism is always a critique of normative function and form that is the grounds for the western theological-philosophical epistemology, including the way this episteme produces race, gender, sexuality, class, ability, nationality as categorical distinctions. Noise—joyful as it is—gets in the way of such smooth, easy conceptualizations. (p. 167)

Noise, which Western thought treats "as always in need of abatement," is vibrational and as such, "begs its being heard, its being listened" (Crawley, 2017, p. 139). Noise, like Blackness itself, bears an excess that resists the colonizing fantasy of compartmentalization and control (Crawley, 2017). Similar to aspects of Thompson's (2017) and Serres's (1982) conceptualization of noise, Crawley flips the script on who and what is considered noise and therefore who and what is considered desirable or disposable. While "noise, in general, became racialized as the other in Europe, as the other of rationality, as the other of the proper" (p. 140), Crawley (2017) focuses on the "joyful noise of Blackpentecostal aesthetics" as a means of "sounding out a way, rehearsing a mode of sonic production that refuses origin and purity, utilizing melismatic, melodic irruption, irreducible noise" (p. 144). The "interruptive" testimonies and joyful tarrying during Blackpentecostal church services are acts of vibrational joy that refuse epistemological demands of quietude, which are associated with "clear" and "pure" thought. "Joyful noise, the noise of Blackpentecostal aesthetics, operates from a different epistemological decentering, a centrifugitive refusal of centeredness" (p. 143). Crawley's reference to "centrifugitive refusal of centeredness" resonates with Indigenous conceptions of refusal (e.g., Grande, 2018; Tuck & Yang, 2014), which seek otherwise possibilities for reconstructing culture and tradition apart from demands for recognition and reconciliation with the state.

Although Crawley's analysis of noise is an indictment of Western theology and philosophy, his analysis can be extended to data science and data science education because of how dominant scripts about data science and statistics align with—and in some instances can be traced directly to—Western theological and philosophical thought (Hacking, 1990). Indeed, "Blackpentecostal aesthetics, to assert again, do not *belong* to Blackpentecostals but can be enacted. Blackpentecostals simply carry the tradition in secreted and clandestine modalities"

(Crawley, 2017, p. 167). In addition to echoing Thompson (2017) and Serres's (1982) claims about the generative force of noise and Jon's (2022) analysis of the productive contributions of marginalized people who have been deemed as noise, Crawley (2017) brings important attention to noise as liberative, embodied, and rooted in Black joy. Crawley considers such joy atheological and aphilosophical insofar as Western theology and philosophy seek closure of openness, while Blackpentacostalism is about "possibilities that exist in plurality for those that have been rejected from the zone of the human" (p. 37). What would data science and data science education look like if it similarly rejected categorization and embraced liberative, embodied, and vibrational joy?

First, a data science and data science education rooted in joyful noise would reject the distinction between the world of sound as "pure reason" and the world of the body and flesh as "impure." The testimonies and tarrying within Blackpentecostal churches is not just noise in a sonic sense; it is also the noise of a body experiencing spiritual excess. As Crawley (2017) writes, "Noise-making from within the Blackpentecostal episteme is choreosonic; it is always about metaphorical and material movement. All sound is motion, all motion movement, all movement choreosonic" (p. 146). D'Ignazio and Klein (2020) implore readers to embrace emotion, embodiment, and viscerality when working with and representing data. In a similar way, data science and data science education could be treated as a choreosonic practice that treats sound/movement and mind/body as co-constitutive. Instead of understanding data exclusively within the domain of "pure reason," choreosonic practices can inspire an analytic approach to data that is not only cognitive but also affective and rooted in the realities of living in bodies that are raced, gendered, classed, and dis/abled according to broader social processes. Further, a choreosonic understanding of data science and data science education reiterates what

scholars such as D'Ignazio and Klein (2020), among many others, have discussed with respect to authentic data-driven investigations: data science is not merely a matter of listening to data (the sonic) but using data to take action for purposes of individual and collective empowerment (movement).

Second, joyful noise is a rupture of Western theological and philosophical assumptions about teleological progress and risk aversion and therefore can be used to rethink widespread narratives about the benefits of a data-driven future. Mikulan and Sinclair (2023) surface dominant assumptions about time in education, such as the assumption of time as arrow and the use of deferred time to promise students that education today will reap benefits in the future. Likewise, Crawley (2017) argues that "Blackness, Blackpentecostal aesthetic practice, disrupts the logic of linear time and space" (p. 146). Crawley does not reject the value of future-oriented thinking. Like Mikulan and Sinclair, Crawley seeks to disrupt and add nuance to future-oriented thinking grounded in notions of progress. He does so by pointing to historical instances of noise—as breath, screams, shouts—as produced and experienced by Black people as they sought, and continue to seek, freedom, liberation, and otherwise possibilities for themselves and their loved ones. Crawley (2017) writes:

> Such centrifugitive time would cause us to rethink the noise in praise houses and hush harbors, the joyful noise of Blackpentecostal practice, because the rupture of liberation and freedom so desired by enslaved peoples then and those of us marginalized through racialization today can now be considered potential time. Potential time because performance practice of blackness is about the certainty, the spiritual conviction, of the to-come liberative possibility, and is the living out of liberation as belief in the flesh. (p. 149)

Although the future is not guaranteed, joyful noise invites a consideration of "potential time" based on "spiritual conviction" and "belief in the flesh." It is an understanding of the future grounded in hope and spiritual certainty as an alternative to the rationalist certainty of the kind

promoted by data-driven, probabilistically-defined scientific practices. Crawley (2017) calls out the tendency of Western philosophy and theology—and I argue by extension, data-driven scientific practices—to focus on "crisis aversion." He states:

> A theology-philosophy of aversion forestalls otherwise possibility, making normative claims on what any set of behaviors, or what any group of people, could and should possibly be. The 'crisis' as the thing so believed to be the only future is but one possibility. To name the thing as crisis averted is to strip away the fact that the infinite set of capacities to be otherwise exists alongside the possible crisis, but it is also to refuse the fact that the infinite other possibilities cannot ever be fully named, claimed, or thought. (p. 177)

Earlier, I had argued that individualistic risk management is a prominent script within dominant approaches to data science education reform. Here, Crawley offers another framing of the matter. Risk management—as an instantiation of the much older notion of crisis aversion rooted in Western philosophy and theology—is a foreclosing of otherwise possibilities through the very act of identifying something as a 'crisis' or a 'risk'. Data science and data science education, therefore, serve the role of lending authority to notions of risk management through the language of probability. Based on this characterization of data science, one might then ask if it is even possible to have a data science or data science education grounded in joyful noise. Or do the assumptions of pure reason, linear temporality, and crisis aversion embedded within data-driven and statistical practices render data science and joyful noise inherently incompatible?

At the very least, joyful noise invites a rethinking and "intra-rupting" of data science and data science education. In her discussion of Serres's transdisciplinary approach to theorizing noise and the parasite, Thompson (2017) states, "The distortion of ideas, models and theories when they are taken outside their disciplinary context is not only a necessary risk but also – more importantly – a possible source of invention" (p. 57). Here, the concept of joyful noise is theorized within Blackpentacostalism, and therefore any attempt to draw out lessons for data

science and data science education necessarily entails transferring the theory of joyful noise into a new context. Crawley's theorization becomes a form of noise that exerts an intra-ruptive force on the disciplinary assumptions of data science and data science education. "The perturbation of these disciplinary 'messages' comes with potentially unexpected insights that may allow alternative ways of understanding phenomena and their operations" (Thompson, 2017, p. 57). Joyful noise is a call to embrace transdisciplinarity within data science and data science education, not only to expand what might be possible through data but also to call into question fundamental epistemological and ontological assumptions about the notion of data itself. In this sense, "noise is both a recurring theme and a strategy of inquiry" (Thompson, 2017, p. 57) because it simultaneously challenges the fetishization of signal over noise and provides a theory for inquiring into otherwise possibilities for data science rooted in attending to moments of rupture, the agentic capacities of the material world, and the perturbing force-relations of people who have been historically dismissed as "mere noise."

Conclusion: Intra-Actions of Signal and Noise

In this paper, I surfaced a selection of scripts and valences in prominent data science education policy documents and initiatives (workforce preparation, individualistic risk management, and surface level calls for access and opportunity for all students) and used them as a starting point to offer critiques of and alternatives for thinking about certain aspects of data science education. In my effort to offer such alternatives, I engaged in the theoretical task of diffractively reading the concepts of agential realism, data science, data science education, and noise through one another. In doing so, I sought to reconsider my use of agential realism in theorizing, re-conceptualize noise in terms of intra-ruptions, and re-examine the onto-ethico-epistemologies of data science and data science education. Such reading raised several questions

about data science and data science education: what is the role of the material world in data science and data science classrooms; what and who is treated as noisy and liminal within data science and to what effect; how do these noisy and liminal agents speak back against and intra-rupt the disciplinary concerns of data science and data science education; and how might data science and data science education be reconceptualized in terms of joyful noise as an alternative to logics of categorization? The purpose of this discussion is not to elevate noise over signal in a way that reproduces a binary between the two ideas. Thompson (2017) states, "Such an understanding of noise thus allows for a dismantling of the hierarchical relationship of signal and noise, first by understanding the relational positions of sender, receiver and noise as interchangeable (the host becomes the parasite and the parasite becomes the host) and second, by recognizing noise as an essential component of material relations (the parasite is constitutive of the relation)" (p. 62). Noise raises attention to its own ambiguity. Noise can be beneficial but it can also be harmful. The contribution of this analysis is to underscore that in either case, noise is a productive force-relation (Thompson, 2017) that gives rise to new possibilities that should not be ignored, particularly in data science and data science education, which can too often be caught up in neoliberal narratives of signal and control.

Reading noise through agential realism through data science provides an alternative to dominant scripts and valences about data science, data science education, and predictions about a future driven by data. In this alternative script, there is value to be found both in traditional data practices and in otherwise approaches to data science and data science education. Following Barad (2007), one cannot judge in advance the ethical implications of data-driven practices. Instead, data-driven practices—like all practices—form cuts that entail attunement to their effects. I do not reject the fact that there may be valid economic concerns around data science or

that data science cannot be used to address concerns such as climate change, unjust policing practices, or economic inequality. However, I argue that scripts about workforce preparation, risk management, and access for all lack a relational understanding of data, let alone a nuanced understanding of the complex interplay between signal and noise within data practices. Diffractive understandings of agential realism and noise promote scripts of data grounded in relations. Relations are always marked with parasites, meaning that they are rarely completely harmonious and often entail patterns of messiness, static, and interruption. This paper promotes the concept of intra-ruption to underscore that who and what is considered noise/static/parasite/the interrupter is itself an agential cut formed within phenomena. I take advantage of this shifting dynamic between signal/information and noise/static/parasite to talk about how noise arises in various ways throughout data science and data science education. This ambiguity is the central point of this paper: rather than understanding the purpose of data science and data science education as the elevation of signal over noise, these fields can alternatively be understood as the intra-active entanglement of signal and noise together.

I conclude this paper by revisiting Crawley's (2017) discussion of noise through the lens of Blackpentacostalism. In a previous section, it was stated that Crawley (2017) describes Blackpentacostal noise as a disruption of logics of space and time. He grounds his argument in particular praise practices that occur during church services, where members engage in choreosonic performances of songs and chants and where "[p]unctuating the chant are hand claps, are the sounds of the bass and snare drum, of the cymbals, of Saints praising noise-like together" (p. 162). Importantly, as Crawley argues, these choreosonic performances are not "interruptions" of the church service in the normal sense:

> When the Saints call Testimony Service doesn't so much begin as much as it happens, as it eventuates, as it anoriginarily opens. Someone might sing a song or

> lead a prayer, but the service doesn't 'begin,' because such a concept would presume that the work of Spirit is in need of being convoked (Crawley, 2017, p. 159)

Interruption implies the existence of linear time. In order for noise to "interrupt" a message, the message must be seen as moving from one beginning to an end, with the noise interceding somewhere in the middle. This is the model of noise that was presented by Shannon and Weaver (1963) and that has come to dominate data-driven and statistical thought. That is, the goal of data science is to help its user move from a state of less information to a state of greater information, with anything in-between that "interrupts" this process serving the role of noise and anything aiding in the production of information serving the role of signal. Thompson (2017) and Serres (1982) challenge this model of dichotomously pitting signal against noise by calling into question who and what counts as signal/host versus noise/parasite. This paper's use of the term "intra-ruption" is intended to capture this ambiguity. Crawley's (2017) analysis of the choreosonic performances within Blackpentacostalism offers another way to understand intra-ruption. Namely, intra-ruption resonates with the temporal disruptions of choreosonic joyful noise, suggesting that the intra-active dynamic between signal and noise within data-driven practices is in reality a dynamic between beginning and end. Normative scripts and valences of data understand data-driven processes as the initiation of an action (hypothesis generation and data collection) and the conclusion of an action (data representation and reflection). Intra-ruptive noise offers an alternative: data intra-rupts and becomes intra-rupted always in the middle of phenomena. Data, in other words, need not be seen as an extractive process with an established beginning and end. Rather, working with data for liberatory ends is a matter of attuning to the ongoing potential of phenomena that are always present. Rather than using data to "capture" aspects of phenomena for the benefit of greater human insight and so-called self-determination,

data science can be a means of dwelling within a world where rich practices of knowing and being already exist. How might data science education look, feel, and sound differently if dominant scripts of signal over noise were replaced with something more nuanced, affective, ambiguous, and intra-ruptive?

REFERENCES

Ahmed, S. (2008). Open forum imaginary prohibitions: Some preliminary remarks on the founding gestures of the 'new materialism'. *European Journal of Women's Studies, 15*(1), 23-39.

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.

Barad, K. (2010). Quantum entanglements and hauntological relations of inheritance: Dis/continuities, spacetime enfoldings, and justice-t-come. *Derrida Today, 3*(2), 240-268.

Barad, K. (2011). Nature's queer performativity. *Qui Parle, 19*(2), 121-158.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D.A. (2020). *PreK-12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education*. American Statistical Association and National Council of Teachers of Mathematics.

Beck, U. (1992). *Risk society: Toward a new modernity*. Sage Publications Inc.

Carroll, S. (2024). TEDxUArizona: Indigenous peoples breathing data back. *YouTube*. Accessed at: https://www.youtube.com/watch?v=jPS_3mZXWXw&t=415s

Collins, P. H. (2000). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. Routledge.

Coyne, K. M. (2020, December 17). White noise, white silence: Who gets to be loud in today's America? *Literary Hub*. Accessed at: https://lithub.com/white-noise-white-silence-who-gets-to-be-loud-in-todays-america/

Crawley, A. T. (2017). *Blackpentecostal breath: The aesthetics of possibility*. Fordham University Press.

Darville, S., & Disare, M. (2016, January 7). In first 'Algebra for All' effort, city will push schools to centralize fifth grade math. *Chalkbeat*. Accessed at: https://ny.chalkbeat.org/2016/1/7/21104576/in-first-algebra-for-all-effort-city-will-push-schools-to-centralize-fifth-grade-math

DataScience4Everyone (n. d.). About the coalition. Accessed at: https://www.datascience4everyone.org/about

Davis, N. (2009). New materialism and feminism's anti-biologism: A response to Sara Ahmed. *European Journal of Women's Studies, 16*(1), 67-80.

de Freitas, E. & Sinclair, N. (2013). New materialist ontologies in mathematics education: the body in/of mathematics. *Educational Studies in Mathematics, 83*(3), 453-470.

Deleuze, G., & Guattari, F. (1987). *A Thousand plateaus: Capitalism and schizophrenia*. University of Minnesota Press.

Derrida, J. (1994). *Specters of Marx: The state of the debt, the work of mourning, and the new international*. Routledge.

D'Ignazio C., & Klein, L. F. (2020). *Data feminism*. The MIT Press.

Dixon-Román, E. (2017). Toward a hauntology on data: On the sociopolitical forces of data assemblages. *Research in Education, 98*(1), 44-58.

Ellenwood, C., & Foxworth, R. (2024). Data back: Indigenous material culture, philanthropy, and Indigenous data sovereignty. *The Public Historian, 46*(2), 123-143.

Fiore-Gartland, B., & Neff, G. (2015). Communication, mediation, and the expectations of data: Data valences across health and wellness communities. *International Journal of Communication, 9*(1), 1466-1484.

Flatley, D. (2021, June 9). Senate overwhelmingly passes $250 billion tech investment bill aimed at countering China. *Time*. Accessed at: https://time.com/6072196/senate-bill-tech-china/

Fotopoulou, A. (2021). Conceptualising critical data literacies for civil society organisations: Agency, care, and social responsibility. *Information, Communication & Society, 24*(11), 1640-1657.

Gaskins, N. (2021). *Techno-vernacular creativity and innovation: Culturally relevant making inside and outside of the classroom*. MIT Press.

Gonzalez, X. (2022, August 1). Why do rich people love quiet? The sound of gentrification is silence. *The Atlantic*. Accessed at: https://www.theatlantic.com/magazine/archive/2022/09/let-brooklyn-be-loud/670600/

Gordon, A. F. (1997). *Ghostly matters: Haunting and the sociological imagination*. University of Minnesota Press.

Grande, S. (2018). Refusing the university. In E. Tuck, & K. W. Yang (Eds.), *Toward what justice? Describing diverse dreams of justice in education* (pp. 47-65). Routledge.

Gutierrez, K., Rymes, B., & Larson, J. (1995). Script, counterscript, and underlife in the classroom: James Brown versus Brown v. Board of education. *Harvard Educational Review, 65*(3), 445-472.

Hacking, I. (1990). *The taming of chance*. Cambridge University Press.

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies, 14*(3), 575-599.

Hardy, L., Dixon, C., & Hsi, S. (2020). From data collectors to data producers: Shifting students' relationship to data. *Journal of the Learning Sciences, 29*(1), 104-126.

Ingold, T. (2011). *Being alive: Essays on movement, knowledge and description*. Routledge.

James, R. (2019). *The sonic episteme: Acoustic resonance, neoliberalism, and biopolitics*. Duke University Press.

Jon, I. (2022). Reimagining the future with liminal agents: Critical interdisciplinary STS as manifestos for anti-essentialist solidarities. *Culture, Theory, and Critique, 63*(2-3), 136-153.

LaMar, T., & Boaler, B. (2021). The importance and emergence of K-12 data science. *Phi Delta Kappan, 103*(1), 49-53.

Lee, V. R., Pimentel, D. R., Bhargava, R., & D'Ignazio, C. (2022). Taking data feminism to school: A synthesis and review of pre-collegiate data science education projects. *British Journal of Educational Technology, 53*(5), 1096-1113.

Lee, V. R., Wilkerson, M. H., & Lanouette, K. (2021). A call for a humanistic stance toward K–12 data science education. *Educational Researcher, 50*(9), 664-672.

Loukissas, Y. A. (2019). *All data are local: Thinking critically in a data-driven society*. MIT Press.

Malaspina, C. (2018). *An epistemology of noise*. Bloomsbury Academic.

Martin, D. B. (2019). Equity, inclusion, and antiblackness in mathematics education. *Race, Ethnicity, and Education, 22*(4), 459-478.

McQuillan, D. (2018). Data science as machinic neoplatonism. *Philosophy & Technology, 31*(2), 253–272.

Mertala, P. (2020). Data (il)literacy education as a hidden curriculum of the datafication of education. *Journal of Media Literacy Education, 12*(3), 30-42.

Mikulan, P., & Sinclair, N. (2023). *Time and education: Time pedagogy against oppression*. Bloomsbury.

National Academies of Sciences, Engineering, and Medicine (NASM) (2018). *Data science for undergraduates: Opportunities and options*. The National Academies Press.

Rosiek, J. L., Snyder, J., & Pratt, S. L. (2020). The new materialisms and Indigenous theories of non-human agency: Making the case for respectful anti-colonial engagement. *Qualitative Inquiry, 26*(3-4), 331-346.

Rubel, L. H., Peralta, L. M., Herbel-Eisenmann, B., Jiang, S., Kahn, J. B., & Lim, V. Y. (2021). Theorizing data science education: An intersectional feminist perspective on data, power,

and “playing the game”. In *Proceedings of the 43rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA)*. Philadelphia, PA.

Sanches, P., Howell, N., Tsaknaki, V., & Helms, K. (2022). Diffraction-in-action: Designerly explorations of agential realism through lived data. *CHI Conference on Human Factors in Computing Systems (April 29-May 5, 2022)*. ACM. New York, NY.

Sellar, S., & Thompson, G. (2016). The becoming-statistic: Information ontologies and computerized adaptive testing in education. *Cultural Studies ↔ Critical Methodologies, 16*(5), 491-501.

Serres, M. (1982). *The parasite*. The Johns Hopkins University Press.

Shannon, C. E., & Weaver, W. (1964). *The mathematical theory of communication*. The University of Illinois Press.

Smith, S. (2002). *Digital signal processing: A practical guide for engineers and scientists*. Newness.

Stornaiuolo, A. (2020). Authoring data stories in a media makerspace: Adolescents developing critical data literacies. *Journal of the Learning Sciences, 29*(1), 81-103.

Thompson, M. (2017). *Beyond unwanted sound: Noise, affect and aesthetic moralism*. Bloomsbury Publishing USA.

Todd, Z. (2016). An Indigenous feminist’s take on the ontological turn: ‘Ontology’ is just another word for colonialism. *Journal of Historical Sociology, 29*(1), 4-21.

Tuck, E., & Yang, K. W. (2014). R-words: Refusing research. In M. W. Winn, D. Paris (Eds.), *Humanizing research: Decolonizing qualitative inquiry with youth and communities* (pp. 223-248). SAGE Publications.

van der Tuin, I. (2008). Response to Sara Ahmed’s ‘Imaginary prohibitions: Some preliminary remarks on the founding gestures of the “new materialism”’. *European Journal of Women’s Studies, 15*(4), 411-416.

Van Loon, J. (2002). *Risk and technological culture: Towards a sociology of virulence*. Routledge.

Van Wart, S., Lanouette, K., & Parikh, T. S. (2020). Scripts and counterscripts in community-based data science: Participatory digital mapping and the pursuit of a third space. *Journal of the Learning Sciences, 29*(1), 127-153.

Walter, M., & Andersen, C. (2013). *Indigenous statistics: A quantitative research methodology*. Walnut Creek.

Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics, 96*, 33-47.

Weiland, T., & Williams, I. (2023). Culturally relevant data in teaching statistics and data science courses. *Journal of Statistics and Data Science Education*, 1-16, https://doi.org/10.1080/26939169.2023.2249969

Wynter, S. (2003). Unsettling the coloniality of being/power/truth/freedom: Towards the human, After Man, its overrepresentation–An argument. *CR: The New Centennial Review, 3*(3), 257-337.

CHAPTER 3: STORYTELLING, MATERIALITY, TEMPORALITIES, AND RISK IN A DATA-INTENSIVE LEARNING ENVIRONMENT

Introduction

I remember talking about how there's such a bigger picture with data and you read, oh, this percent of something. And it's like, okay, well that's like a cold, hard fact, but really no. It's like, who surveyed those people? Who were the people surveyed? What… why was… who were the people surveyed? There's so much story behind every single piece of data and you can't even fully take the data to be, oh, this is a fact, without really understanding where it comes from and who it comes from and who the data's about.

The quote above was spoken by Adele, an undergraduate first-year student enrolled in an interdisciplinary course on data storytelling in a large public university in the U.S. Midwest. Her words gesture toward a question that this article seeks to address: what ontological and epistemological orientations toward data emerged within the data storytelling course, and what are the implications for instructional design and pedagogy within data-intensive learning environments?

In recent years, data have been described as "the new oil" (Couldry & Mejias, 2019, p. 1). The perceived importance of data in society, along with the promises of profit suggested by the metaphor of data as oil, has given rise to numerous calls to expand educational offerings in data science and data literacies throughout K-16 curricula in the United States (e.g., Bargagliotti et al., 2020; Data Science for Everyone, n.d.; Draper, 2020). Efforts to promote the teaching and learning of data have been most evident at the post-secondary level, where universities, bootcamps, and training programs have proliferated to meet the surge in demand for employees with data literacies and data science experience across various fields (Draper, 2020).

The terms data literacies and data science do not have a settled definition but broadly include the ability to collect, make sense of, and represent data. Data acumen encompasses not only familiarity with the quantitative and computational tools necessary to make sense of data

but also other interdisciplinary skills such as the ability to decode information and make connections between data representations and real-world scenarios (Bargagliotti et al., 2020). A rapidly growing group of scholars, however, argues that data acumen should include more than technical data skills. Instead, they call for supplementing technical proficiency with knowledge of and a desire to act with respect to issues of social justice, equity, and broader sociopolitical and cultural concerns related to data and data practices (D'Ignazio & Klein, 2020; Fotopoulou, 2021; Lee et al., 2021; Lim et al., 2022; Philip et al., 2016). Students' learning interactions with data "are far more complex and wide reaching than are often presented in curricula and professional development materials" (Lee et al., 2021, p. 1). It follows, then, that data literacies and data science education efforts should seek more expansive conceptualizations of what it means to teach and learn about data. This study is one attempt to contribute to such efforts.

This paper asks what possibilities and provocations emerged within, through, and from a course on data storytelling when one "thinks with" (Jackson & Mazzei, 2011) concepts from feminist new materialisms and science technology studies (Barad, 2007), anthropology (Ingold, 2011), Indigenous wisdoms (Cajete, 2000), and educational philosophy (Bingham & Sidorkin, 2004). Moreover, this article considers the implications of such possibilities and provocations for instructional design practices within data-intensive learning environments. I conceive of instructional design as the iterative process of constructing, carrying out, evaluating, and adjusting instruction, along with the theories about students, classrooms, and the sociopolitical context of teaching and learning that guide this work. Many instructional design models rely on overly humanist, rationalist, and individualistic accounts of teaching and learning, as exemplified by the following statement on instructional design: "Learning is a personal and covert cognitive activity, which is idiosyncratic to an individual. Individuals who construct knowledge and skills

accomplish learning" (Branch & Kopcha, 2014, p. 79). One of the most prevalent models for instructional design is the ADDIE (analysis, design, development, implementation, and evaluation) model, which is a "generic systems approach, similar to that applied in other fields such as software engineering and product design" (Molenda et al., 2006, p. 576). In this view, instructional designers treat classrooms as similar to laboratories or corporate product development spaces where interventions can be designed to manage teacher and student activity in a scientific manner.

I put forth an alternative vision for instructional design and pedagogy in the context of data literacies and data science education. Some instructional design models incorporate attention to the broader learning environment (Molenda et al., 2006), such as problem-based learning and constructivist learning frameworks. Relatedly, some critical scholars promote approaches to instructional and research design and pedagogy that draw on sociocultural, ecological, and decolonial theories of learning, among other theories and perspectives (e.g., Bang, 2017; Bang & Vossoughi, 2016; Vakil et al., 2016; Vossoughi, & Gutiérrez, 2017). This paper seeks to build on these efforts by bringing attention to the material-discursive environment of the data literacies and data science classroom. By attending to relations of materials and discourse, instructional design and pedagogical practices can go beyond solely attending to the rationalist dimensions of teaching and learning and instead attune to the affective dimensions of classrooms. One contribution of this article, therefore, is to shift conversations about instructional design and pedagogy toward an emphasis on materiality, affect, and relationality. Importantly, while this paper is associated with a study where my participants and I sought otherwise possibilities (Green, 2020) for data literacies and data science education, I do not intend to promote binary categorizations that equate traditional approaches to data, instructional

design, and pedagogy with something "bad" and that equate non-traditional approaches to these topics with something inherently "good." Instead, a significant finding of this paper is that as part of my attempt to engage students with thinking about data in novel, critical, and creative ways, traditional understandings about data continued to emerge within the discursive-material environment of the classroom, for better or worse. Instead of categorizing particular conceptualizations of or approaches to teaching data as "good" or "bad," this paper advocates for attuning to the epistemological and ontological orientations to data that emerge within classrooms through close attention to students' words and actions, along with the agentic capacities of the more-than-human world that are always a part of teaching and learning. Paying attention to these cuts (Barad, 2007) entails a responsibility to engage with or disengage from particular ideas about data based on the needs and interests of those involved within the data-intensive learning environment.

This article proceeds in five parts. First, I provide a brief overview of literature relevant for this study. Second, I discuss this study's conceptual framework, which relies on a relational ontology that offers an alternative to the cognitivist and behaviorist accounts of teaching and learning prevalent within the study of instructional design and pedagogy. Third, I discuss my methodological approach and the data generated throughout this study. Fourth, I engage in an extensive discussion of the interdisciplinary course on data storytelling on which this study is based, along with vignettes and analyses of the onto-epistemological ideas about data that emerged throughout the course. Lastly, I conclude by offering implications for data literacies and data science teaching and learning.

Relevant Literature

In previous writing, I suggest that the increasing importance of data throughout society, combined with efforts to expand data science offerings across K-16 settings, motivate an understanding of the epistemological and ontological orientations to data that emerge within data-intensive learning environments. Such understanding provides an opening to consider otherwise possibilities for data science beyond narrow conceptualizations that associate data with objective truth. Understandings of data as mere numbers, facts, and graphs that are divorced from affect, emotion, and personal experience have the potential to exclude students from engagements with data science, especially those students whose values do not align with the overly rationalist worldviews too often promoted within data-intensive learning environments (see Kahn et al., 2022; Lim et al., 2022). This argument mirrors work done in mathematics education, where scholars have argued that mathematics' close ties to Eurocentric rationalist thought decenters non-white, non-male students because Eurocentric thought instead centers "elitism and social stratification that looks to build the economic power of corporate entities and its White male leadership" (Tate, 1995, p. 168, citing Cohen, 1982; Ernest, 1991; Jefferson, 1954; Smith, 1937). In a similar vein, because "the hype around big data and AI is deafeningly male and white and technoheroic" (D'Ignazio & Klein, 2020, p. 9), this can have the effect of isolating students who wish to think about and use data in ways that emphasize issues of power, ethics, and the body.

There are scholars whose work provides a strong foundation for alternatives to data science and data science education that go beyond narrow, rationalist, and "technoheroic" approaches to data. Wilkerson and Laina (2018), whose work is particularly relevant for this paper's study, promote storytelling with and about data as a "fruitful way to build coherence

across multiple statistical contexts" while also highlighting data practices "as a messy, human endeavor" (p. 1224). In their work with middle school students, they asked students to repurpose public data to explore questions about their local communities. One set of data related to a rodent problem that affected their local community, and the other set involved demographic and other municipal data (e.g., income, public school enrollment, and housing) provided on the city's website. The students then discussed their findings with one another through presentations, visual displays, text, and posters. While this paper's study, which is connected to a course on data storytelling with undergraduate students, resonates strongly with the work of Wilkerson and Laina (2018), both studies differ in some respects. Wilkerson and Laina did not give explicit instructions to students to tell stories with data. Instead, they interpreted the students' activities as a form of storytelling "because they were expected to make sense of repurposed data in terms of their own knowledge and experiences, and to construct public artifacts to share the results of their data investigations" (p. 125). In contrast, this study explicitly engaged students with the task of creating various data-driven "stories," where the term story was broadly constructed to encompass any structured form of communication and expression. This study's data storytelling course emphasized—and at times, risked conflating—the relationship between stories and narratives, with the latter described in terms of Freytag's Pyramid (exposition, rising action, climax, falling action, denouement, inciting incident, resolution) (Feigenbaum & Alamalhodaei, 2020). Moreover, although Wilkerson and Laina (2018) examined how students made sense of data and, at many times, saw themselves within the data, this study seeks to more closely attend to deep epistemological and ontological questions about data and how various orientations to data emerge through the discourses and materiality of storytelling practices. The purpose of this explicit focus on storytelling is to more fully embrace the premise set forth in Wilkerson and

Laina that storytelling is a generative concept with which to think about and reimagine data science education, namely building on this work by using the epistemological and ontological orientations to data that emerge within classroom practice as a starting point.

Researchers have sought to understand students' and teachers' general and domain-specific epistemological beliefs and their impact on teaching and learning through the use of teaching scenarios, questionnaires, and self-report measures (e.g., Gill et al., 2004; Hashweh, 1996; Tsai, 2000). Although there may be reasons for studying students' epistemological beliefs from a cognitivist lens, such positivist and post-positivist research aligns with the epistemological and ontological assumptions that underlie the very approaches to data literacies and data science for which this study seeks otherwise possibilities. Instead, this study more closely follows the work of scholars such as de Freitas and Sinclair (2013, 2014), who consider the material and ontological possibilities that emerge within mathematics classrooms, which one does not "uncover" so much as honor through close attention to fine details and a consideration of one's own involvement in research. Part of honoring these possibilities is rethinking the role that human as well as more-than-human bodies play in the "'dance of agency' that makes up mathematical activity" (p. 454, citing Pickering, 1995), as well as rethinking the very boundaries that we demarcate between human and more-than-human bodies. This work is not only a distinctive methodological orientation to qualitative research but also a stance on the nature of mathematics itself, which de Freitas and Sinclair (2013) understand as a "body" that is "an assemblage of human and more-than-human mathematical concepts" (p. 453). Mathematics as an assemblage means that it is primarily relational. In particular, it is in relation with the gesturing and moving bodies that co-constitute it, such that the assemblages of mathematics are "porous and partially closed differential systems that are never entirely cut off from the chaos (pure

difference) that surrounds them, but they are differentiated from it by degree, as though they were contractions of intensity or energy" (de Freitas & Sinclair, 2013, p. 458). In this account, the materiality of mathematics is not about curating manipulatives and studying their affordances and constraints, as if they were mere passive objects ripe for human manipulation and control. Nor is it the case that the materiality of mathematics exists independent of its relation to other bodies. Instead, gestures, diagrams, movements, and matter participate in the "rhythmic engagement" (de Freitas & Sinclair, 2014, p. 155) that constitutes mathematics as a dynamic and ongoing phenomenon.

This study treats the concepts of data, data literacies, data science, and teaching and learning about data in a similar way—that is, in terms of dynamic and contingent phenomena constituted by relations that are not understood through the language of cognition and mediation. This paper emphasizes the materiality of data by understanding data as an assemblage (Dixon-Román, 2017) that invites us to consider the epistemology, ontology, and ethics of data together (Barad, 2007). The next section discusses this paper's conceptual framework further.

Conceptual Framework

I conceive of data and classrooms, as well as teaching and research, in terms of a relational ontology (Barad, 2007; Bingham & Sidorkin, 2004; Cajete, 2000; Ingold, 2011). Feminist new materialisms emphasize (1) seeing human and more-than-human bodies and social and abstract entities as relationally defined, (2) replacing individual agency with flows of affect, which refers to the capacity for one body to affect another, and (3) understanding assemblages as being constantly territorialized and de-territorialized by affective flows (Fox & Alldred, 2014). Relations are dynamic, ever-shifting, and contingent. Massey (2005) states, "These are not the relations of a coherent, closed system within which, as they say, everything is (already) related to

everything else. Space can never be that completed simultaneity in which all interconnections have been established, and in which everywhere is already linked with everything else" (pp. 11-12). Ingold (2011) draws a nuanced distinction between relations that are *between* actors and relations that run *along* them. The former idea ("between") conceives of a world composed of separable beings who act as nodes within a larger connected network of lines that connect these nodes. The latter idea ("along") resonates with Barad's (2007) notion of intra-action (as opposed to interaction), which is the recognition that distinction does not precede but rather emerges from and through encounters. "Distinct agencies are only distinct in a relational, not an absolute sense, that is, agencies are only distinct in relation to their mutual entanglement" (Barad, 2010, p. 267). In a previous manuscript, I diffractively read the ideas of intra-action, agential realism, and noise (Crawley, 2017; Serres, 1982; Thompson, 2017) through one another to suggest that every intra-action entails "intra-ruption." Intra-ruption refers to the notion that all relations involve ruptures, tears, and degrees of messiness caused by what Serres (1982) calls "parasites," which he uses in three ways: as the act of one organism feeding off another, as the perception of an undesirable person in society, and as informational noise. Intra-ruption borrows from the portmanteau of intra-action to suggest that who or what exactly acts in this role of the disruptive parasite is never settled but instead emerged as a co-constitutive element of any phenomena. It is this notion of relations that this study foregrounds.

I also draw on conceptualizations of relational ontologies that avoid placing an exclusive emphasis on relations among humans or on relations solely about human activity. Relations involve both humans and the more-than-human world, and both are intertwined (Cajete, 2000). Relational ontologies are fundamentally material as well as discursive (Barad, 2007). Matter, furthermore, does not refer to inert substance. It is "no longer imagined here as a massive,

opaque plenitude but is recognized instead as indeterminate, constantly forming and reforming in unexpected ways. One could conclude, accordingly, that 'matter becomes' rather than that 'matter is'" (Coole & Frost, 2010, p. 10). Data, in particular, are not abstract numbers that refer to an independently existing reality. Instead, data are "cultural artifacts created by people, and their dutiful machines, at a time, in a place, and with the instruments at hand for audiences that are conditioned to receive them" (Loukissas, 2019, pp. 1-2). Loukissas's (2019) definition of data is an important intervention into resisting the "god trick" of data (D'Ignazio & Klein, 2020) and its hold on our collective imaginations. Much like other perspectives of data within critical data studies, however, this critical view of data is largely humanist because it places humans at the center of data practices, with "dutiful machines" serving as passive mediational devices. Dixon-Román (2017), drawing on Kitchin (2014), goes further by conceptualizing data as an assemblage of discursive-material relations. Data, according to him, are assemblages that are "formed, shaped, and entwined with the histor(icit)y of the human" (p. 50). These assemblages include, for instance, systems of thought, forms of knowledge, materialities and infrastructures, organizations and institutions, the marketplace, and sociopolitical relations. I build on these ideas about the materiality of data by framing the material world as a form of "noise" that traditional approaches to data science and data science education too often seek to minimize, control, or position as mere mediating agents. One purpose of this paper, then, is to speak back against such treatments of the material world as mere noise.

Relational ontologies provide an invaluable opportunity to understand data literacies and data science classrooms in new ways. For instance, "[I]ncreased knowledge about the material and relational complexity of teaching practices can, for instance, be of particular value for the education and support of new teachers" (Klykken, 2021, p. 2). Relational ontologies can enable

researchers and practitioners to attune to the material-discursive environment of the classroom and how this environment opens up or forecloses instructional possibilities from emerging. In discussing the nature of education, Bingham and Sidorkin (2004) state, "It all depends on relations" (p. 1). They state, "A theory of education is, in other words, a theory about the educational relationship. It is not about the 'constituents' of this relationship (i.e., the teacher and the learner) but about the 'relationality' of the relationship" (p. 13). Relations—as part of the intra-active becoming of humans and more-than-humans within classrooms—are a particularly apt area of focus for data-intensive learning environments due to the increasing connections between data and digital and mobile technologies, laws and institutions, and social relations. As Gorur and colleagues (2019) argue, "Studies in education need to take into greater account the ways that teachers, for example, ally with things and technologies of different kinds in their work. These alliances are never innocent, neutral or predictable activities" (pp. 4-5). Paying attention to these alliances with "things and technologies" is a political act. Relational ontologies are a deliberate step away from humanist frameworks that dominate education research and practice.

Methodology

This study uses an intra-active analysis (Barad, 2007; Ehret et al., 2016), with special attention paid to "intra-ruptions," to consider how human and more-than-human bodies and affective intensities within the data storytelling course enacted boundaries and exclusions of concept formation that reproduced or disrupted dominant onto-epistemic orientations around data. Such analysis mobilizes theories from feminist new materialisms, anthropology, and science and technology studies as tools to "think with" (Jackson & Mazzei, 2011) the data that was a part of this study. This data consisted of detailed field notes that I wrote after each class

session, individual and focal group interviews, video recordings of classwork, recorded Zoom sessions, and student work. The student work included survey questions and survey development materials, data story explorations, data postcards, research journals, research posters, a data story artifact, and various other assignments.

My intra-active analytical approach consisted of iterative cycles of abductive inquiry followed by a diffractive (Barad, 2007) reading of this study's data. An abductive approach to data is one where a researcher embraces moments of "breakdown, surprise, bewilderment, or wonder" with respect to their engagements with data (Brinkmann, 2014, p. 722). It is an approach to knowledge production that "presents research as part of the life process, as what we do in situations of breakdown that inevitably arise in life's situations—big or small" (Brinkmann, 2014, p. 722). Brinkman (2014) uses the term "stumble data" to liken research to an experience of stumbling: "This is what a breakdown is: An experience of stumbling, which causes a situation (in the pragmatist sense), and where inquiry is meant to result in a regaining of one's balance" (p. 724). By abductively reading the data, I sought to open myself up to surprise and to encounter insights, connections, and questions that I had not previously considered. Part of this abductive approach to reading the data was using "writing as a method of inquiry" (Richardson & St Pierre, 2005). Following each session of the course on which this study is based, I wrote field notes to myself that discussed what occurred that day; the feelings that I felt before, during, and after class; surprising insights and connections that my students and/or I gained that may be relevant to this study's research questions about the ontological and epistemological orientations toward data that emerged within the course; and insights, connections, and questions that were raised for me about being a teacher, researcher, and graduate student in this space.

In keeping with the embrace of this abductive notion of surprise and bewilderment, I did not limit taking field notes only after each course session. Instead, taking advantage of the portability of computer technology, I was able to write field notes at nearly any place and time. Indeed, particular moments in my day would cause me to stumble on an idea, allowing me to stop the activity that I was currently engaged in and to add new insights into my notes. In doing so, I sought to embrace and sensitize myself toward the way that knowledge production occurs in relation to the contingent, dynamic, and material configuration of the environment, which includes computers, servers, algorithms, and code, along with the inspiration that one can draw from everyday activity. This process of writing enabled me to consider my own involvement in the research process, not as a fully formed researcher examining the classroom from afar but rather as a co-constituted subject-in-formation that underwent a process of becoming just as much as my students, my co-instructor, and the other agencies of the classroom. For instance, as I continued to write into my field notes, I observed that many of my written notes were shaped by events or experiences that were happening at the moment. In one note, I discussed a tension that arose in the class about how much to adjust the course based on student interests. How much of the course should be shaped by my instructional goals and supposed "expertise" as an emerging scholar, and how much should I "go with the flow" based on what students were expressing in class? These tensions emerged because, as I had written in my field notes, I had recently listened to a podcast that inspired this line of thinking. Indeed, my field notes are filled with moments where current life experiences were explicitly informing how I was interpreting the unfolding of the course and how I was adjusting my design and instruction of the course based on these interpretations.

The other part of the iterative cycle of intra-active analysis consisted of a diffractive (Barad, 2007) reading of this study's data. Diffraction is a methodology of "carefully reading for differences that matter in their fine details, together with the recognition that there is intrinsic to this analysis an ethics that is not predicated on externality but rather entanglement" (van der Tuin & Dolphijn, 2012, p. 50, interview with Karen Barad). The diffractive reading of this study's data consisted of attending to the discursive-material moments that emerged throughout the data storytelling course, including the words, gestures, movements, and affective flows that occurred on account of the capacities not only of the people in the classroom but also the more-than-human world of which we are a part. For instance, as I discuss in a later part of this paper, I considered differences in what materials students used to create various data physicalizations and what impact these materials had on the epistemological and ontological orientations to data that emerged throughout the course. However, rather than conceptualizing differences in materials as a mere matter of the affordances and constraints offered to human actors, I conceptualized the materials themselves as possessing their own agentic capacity. This enabled me to understand the differences between these materials—for instance, between a set of Legos, a cookie, and a ball of yarn—as giving rise to various epistemologies and ontologies of data. These differences, in other words, were "differences that matter," insofar as they had their own forceful impact on the world beyond merely serving as a site of comparing and contrasting the affordances and constraints of mediating devices that exist solely for purposes of human manipulation and use. "Differences that matter," moreover, also included broader differences among orientations to data, and how these differences occurred across space and time. For instance, a key finding that emerges in this paper is based on a consideration of how the same student intra-acted with her data physicalization in two very different ways. Instead of labeling one moment as more favorable or

indicative of learning as the other, I attempted to read both moments as differences that produce an effect—namely the creation of a traditional orientation to data and the creation of a corresponding and co-constitutively defined non-representational approach to data. In other words, the term "matter" in the phrase "differences that matter" refers to both matter as a noun (as in, physical substance) and matter as a verb (as in, to be of or give rise to importance or significance). In my approach of using diffraction as a methodology, I sought to consider how differences among moments in my data "mattered" in both senses of the word.

To help organize and write my diffracting reading of the data, I used the software MAXQDA. This software was an agentic participant in shaping how I configured and considered this paper's study and therefore could be considered an aspect of the diffractive measuring apparatus used to read different moments in the data through one another. Importantly, while MAXQDA is conventionally used to code qualitative data, I used a diffractive methodology in lieu of coding. Some researchers caution against coding practices, which too often create categories that "supposedly possess coherent essences and consistent traits for theme-building and subsequent meaning making" (Jackson, 2013, p. 742). This practice "essentializes people and their experiences" (Jackson, 2013, p. 742), leading us to believe that knowledge is a matter of "fitting isolated particulars encountered here and there into categorical frameworks of ever wider generality" (Ingold, 2011, p. 160). The purpose of this study is to attend to the epistemological and ontological orientations that emerged within this course, not for purposes of creating generalities for classroom practice but rather to create resonances (Tracy, 2010). In addition to reducing and essentializing people and their experiences, coding is "too often divorced from the theory that supposedly guides the study" (Augustine, 2014, p. 748). In an effort not to divorce theory from the study and, instead, take seriously the notion that theoretical

concepts are specific physical arrangements that are part of the materiality of the world, this study also uses as a diffractive apparatus of measurement Jackson and Mazzei's (2011) notion of "thinking with theory." It is a non-positivist approach to research—strongly influenced by Barad's (2007) notions of intra-action and diffraction—that emphasizes bringing texts in relation to each other. To think with theory means to go beyond conventional and interpretive approaches to qualitative inquiry and instead, borrowing the words of Deleuze and Guattari (1987), to plug texts into one another towards the creation of something new. Jackson and Mazzei (2011) state, "[W]e are *doing* and *using* the vocabulary and concepts [of different philosophies] as we push research and data and theory to its exhaustion in order to produce knowledge differently; in this way, we focus on the constitutive and generative aspects of texts" (p. 7). Thinking with theory stresses how texts function in their production of truths rather than what preexisting truths exist within a text. "The analyst's job is not to uncover the hidden significance of the work but to describe its constituent parts and their operation" (Bogue, 2003, p. 60). Bringing texts in relation to each other is never settled; it is a continual reworking of the discursive-material configurations of texts, a "constant, continuous process of making and unmaking" (Jackson & Mazzei, 2011, p. 1). These analytic strategies—abduction, concept coding, and thinking with theory—served as "sensitizing devices" (Decuypere, 2019, p. 137) that opened up possibilities for tracing the relational formations and affective movements of the data storytelling classroom.

In particular, instead of loading all of my available data into MAXQDA, I selected a subset of my data based on an abductive orientation to my research process, which was in turn shaped by my field notes. For this paper, this subset of data included the interviews, written assignments, data postcards, and survey-based project materials of three students in the course (Bobble, Saba, and Daisy). I chose this subset of data because the students' intra-active

entanglements with various materials, along with their comments about data, highly resonated with this paper's research question. I soon decided, however, to focus primarily on Bobble and Saba because a diffractive reading of two students' engagements in the course resulted in a more compelling analysis and response to the research question than a reading of all three students' engagements. Such choice is an agential cut (Barad, 2007) that results in inclusions and exclusions, and the choice of focusing on only two of the students at the exclusion of others remains an ongoing tension in this study.

In addition to the images, transcripts, and assignments that I uploaded to MAXQDA, I also kept annotated notes from the relevant literature in a separate document. Following a diffractive approach to reading data, I set the MAXQDA and literature document side-by-side in order to read both through one another. I asked myself: how were my readings of the literature on relational ontologies informing how I understood the activities of Bobble, Saba, and their surrounding materials? How did their engagements shape how I was understanding, agreeing with, or disagreeing with the ideas about data promoted within the literature? MAXQDA enabled me to take notes directly on the images in the dataset and to add comments directly onto the text. Consistent with this paper's research question, these notes primarily dealt with how various moments in the dataset—as shaped by and shaping my reading of the literature—gave rise to various epistemological and ontological orientations to data. As I wrote these notes, I grappled with two primary tensions. The first tension involved my dual role as an instructor and researcher for the course. Many of Bobble and Saba's remarks about and engagements with data did not correspond with what I had hoped the students would learn or think about data. This dissonance surprised me and led me to change the overall direction and scope of this paper. It also forced me to confront my own desires as a teacher and researcher—should the goal of this project have

been to get students to conform to my theoretical orientations toward data, or should the goal of this project have been to support students in developing their own unique understanding of data, even if it departs from my own philosophical commitments and, in some cases, reproduces potentially harmful epistemological and ontological orientations to knowledge production and use? This tension arose in my diffracting reading of the data, as I felt an urge to dismiss or set aside moments in the data that did not resonate with my own worldview. The second tension involved the impact of MAXQDA, my research process, and my tendencies and patterns of thought as an emerging scholar. MAXQDA enables users to add comments directly onto text and, most importantly, affix a label onto the comments. Moreover, my previous scholarly training emphasized deriving general themes from text and locating patterns within large amounts of data, this training meant that I was inadvertently coding my data despite my desire to avoid coding for the reasons that I described above. I do not claim to have resolved these tensions in this manuscript. Instead, consistent with my theoretical commitments to noise, ambiguity, and disruption, I choose to sit with these tensions and understand my enactment of a diffractive reading of my data as necessarily "impure" and marked by tendencies—shaped by my material environment and my previously scholarly training—to read data in more conventional ways. My intra-active analytical approach, therefore, is an ongoing development rather than "representative" of a mastery over a particular methodological technique. I believe that attuning to such imperfection is nevertheless aligned with an intra-active orientation to knowledge production, which refuses to pretend that research occurs in a vacuum divorced from the forces that shape one's research.

Relational Practices, Affective Intensities, and Flows of Discourse-Materials: Onto-Epistemic Orientations to Data Emerging in a Course on Data Storytelling

I begin this section with a description of the data storytelling course in which occurred the relational practices, affective intensities, and flows of discourse-materials that are the subject of this article's socio-material intra-active analysis. This description provides an overview of the course, including a brief explanation of its origins, the planning process, the intended goals of the course, the students who enrolled in the course, and a selection of their work throughout the course. The purpose of this description is to provide the contours from which onto-epistemic assumptions and new possibilities for data emerged. The course was the"context" for the study, but I do not intend to use the word context in a way that suggests a passivity on the part of the course. On the contrary, my analysis seeks to show how the course—including the classroom, the conference table where students did their work, the course assignments, and myself—was an agentic participant in the movements and circulations that gave rise to teaching, learning, and meaning-making in relation to data.

Following the description of the data storytelling course, I organize my analysis along two main sections. In the first section, I describe a vignette that focuses on the activities of two students, Saba and Bobble, as they presented and discussed Lego constructions as part of a data postcard activity. Then, I describe how this vignette highlights the materiality of data by drawing attention to how gestures, words, and the agentic capacities of discourse-materials within the classroom space enacted particular boundaries and exclusions about data, thereby opening up some possibilities for learning while closing others. In the second section, I present a vignette of Saba, Bobble, and Daisy's year-long research project for the course, which involved surveying students about their attitudes on COVID-19 masking. The vignette describes the yarn-based data

physicalization that they designed as part of the project and that they presented at an undergraduate research conference. I then describe how this vignette, along with the previous vignette, highlight data's temporal dimensions. This temporality emerged from the intra-activity among the students' words and actions, the materiality of their data physicalization, and broader discourses about time, uncertainty, and risk as they connect to data and statistics.

Data Storytelling Course

The data storytelling course emerged from a two year long collaboration between me and a professor in sociology who teaches in a residential living and learning subdivision of the university focused on public affairs. The professor and I shared an emerging interest in incorporating data literacies skills into the undergraduate curriculum. Before teaching the data storytelling course, we co-developed and carried out a pilot project that took place in a different course on public affairs. In that project, we studied how students thought about data relevant to contemporary social issues. For instance, as part of the study, we asked students to read a New York Times article that contained graphs and maps showing that wealthier and whiter neighborhoods have healthier and fuller tree canopies. The students discussed how they related or did not relate to the graphs and maps and how they might reimagine the future based on them. Drawing on these experiences, we designed and proposed the data storytelling course through our university's Honors College, which provided funding and a space to create a course from the ground up.

The course was a research seminar that engaged students with various aspects of data storytelling. There were nine students in the course. Their demographic information and pseudonyms are displayed in Table 1.

Table 1. Students in the Data Storytelling Course

Name	**Gender Identification**	**Race/Ethnicity Identification**	**Year**	**Known Major/Area of Study**
Constance	Female	White	1st	Psychology
Dawn	Female	White	1st	Business
Adele	Female	White	1st	Computer Science
Eve	Female	White	1st	Communication Arts
Saba	Female	South Asian	1st	Pre-Medicine
Frodo	Female	White	1st	International Relations
Bobble	Female	White	2nd	Public Relations
Sasha	Female	Middle Eastern / White	1st	English/Psychology
Daisy	Female	Asian	2nd	Pre-Medicine

The course occurred in Fall 2022 with the expectation that coursework would continue in Spring 2023. Based on requirements from the Honors College, the course culminated in a research experience where students presented findings at an undergraduate research and arts forum in April 2023. Following various discussions about our shared and unique research interests related to data and the goals we had for the course, the professor and I decided in the Summer 2022 that the research experience would consist of two parts: in groups of three, the students would (1) choose a social issue and design a survey that asked other undergraduate students to respond to questions about data visualizations related to that social issue, and (2) create a data-driven story based on the survey results. The students and instructors developed the following research questions for the project: (1) How do undergraduate students respond to data artifacts and questions connected to contemporary social issues dealing with food insecurity, healthcare, and

economic mobility? (2) How might the creation of critical and creative data-driven stories provide insights into how students make sense of data from an interdisciplinary perspective?

In the Fall 2022, we held 12 in-person class sessions for 1 hour and 20 minutes each. I took the lead in designing and teaching each class session, with the professor providing guidance and shaping the overall nature of the course. The required textbooks were Data Feminism (D'Ignazio & Klein, 2020), The Truthful Art: Data, Charts, and Maps for Communication (Cairo, 2016), and Designing and Doing Survey Research (Andres, 2012). The course also included readings about various definitions of data, various approaches to storytelling and data storytelling, and issues of power and ethics around data. During the second half of the semester, I began to share theories that I was learning as a becoming teacher-researcher, including the idea of agential cuts (Barad, 2007), the agency of more-than-humans, the inseparability of the observers and observed in research, the situatedness of data and knowledge (D'Ignazio & Klein, 2020; Haraway, 1988), and the hauntologies of data assemblages (Dixon-Román, 2017). In addition to the survey project, we assigned three "data story explorations," where we asked students to write about a data story of their choice. The purpose was to develop students' familiarity with and appreciation for the many different kinds of data stories that emerged due to technological developments and advancements in the data visualization community, the media, academia, and various artistic communities.

The first two months of the course consisted of discussions about the nature of data and data storytelling, the role of power in data practices, and technical details of survey design. During this time, everyone in the course collaborated to design a survey related to the topics of food insecurity, healthcare, and economic mobility. Each three-person group chose these topics at the start of the semester. Each group designed one page of survey questions based on a few

specifications: the questions should incorporate some kind of graph, table, or other data form; the survey questions should contain a mix of open-ended and closed-ended questions; and the survey questions should be responsive to the research questions posed above. Using the collaborative design features in the Qualtrics survey software, the class combined all the survey questions into a single survey, along with the required consent language, language pertaining to compensation for survey participants, and questions that solicited demographic information from survey participants. Everyone also collaborated to produce all the materials required for IRB approval of the survey project. Finally, we consulted various deans within the university prior to IRB approval, which we obtained in November 2022.

The last two months of the course consisted of various discussions and activities about data storytelling. The original intent was to spend these last two months analyzing survey results and designing data stories. However, several bureaucratic obstacles delayed the dissemination of the survey until the first week of December. For instance, because of a controversial and racially insensitive survey that a professor at the university disseminated in 2019, the university made it more difficult to survey undergraduate students. In addition, everyone in the course wanted to create a raffle for survey participants using $3,000 in funds that had been allocated to the course. After numerous email communications with the university's lawyers, however, the lawyers indicated that an online raffle would not be permitted because it fell under the scope of state gambling laws.

Despite these frustrations, the unexpected delay to disseminate the survey proved fortuitous. In addition to having various lessons on narrative structure and basic statistics, I decided to use the extra time to ask students to create "data postcards." Data postcards are an activity inspired by Lupi and Posavec's (2016) work in Dear Data, a year-long project in which

two information designers created and exchanged hand-drawn postcards based on particular types of data about their lives. In the data postcards activity, we asked the students to create a data postcard based on the number of times they picked up their phones and the reasons they did so, or based on any other data they wanted to collect about themselves. The students then wrote about what the process was like, what the process told them about their lives, how they went about creating their data postcards, and how the process made them feel. Following their presentations of their data postcards, we asked the students to create a second data postcard using the same data and explaining the story of their data. Based on feedback from students, the data postcard activity was formative and eventually led to a final activity in the course where students created a third iteration of their data postcards using Legos.

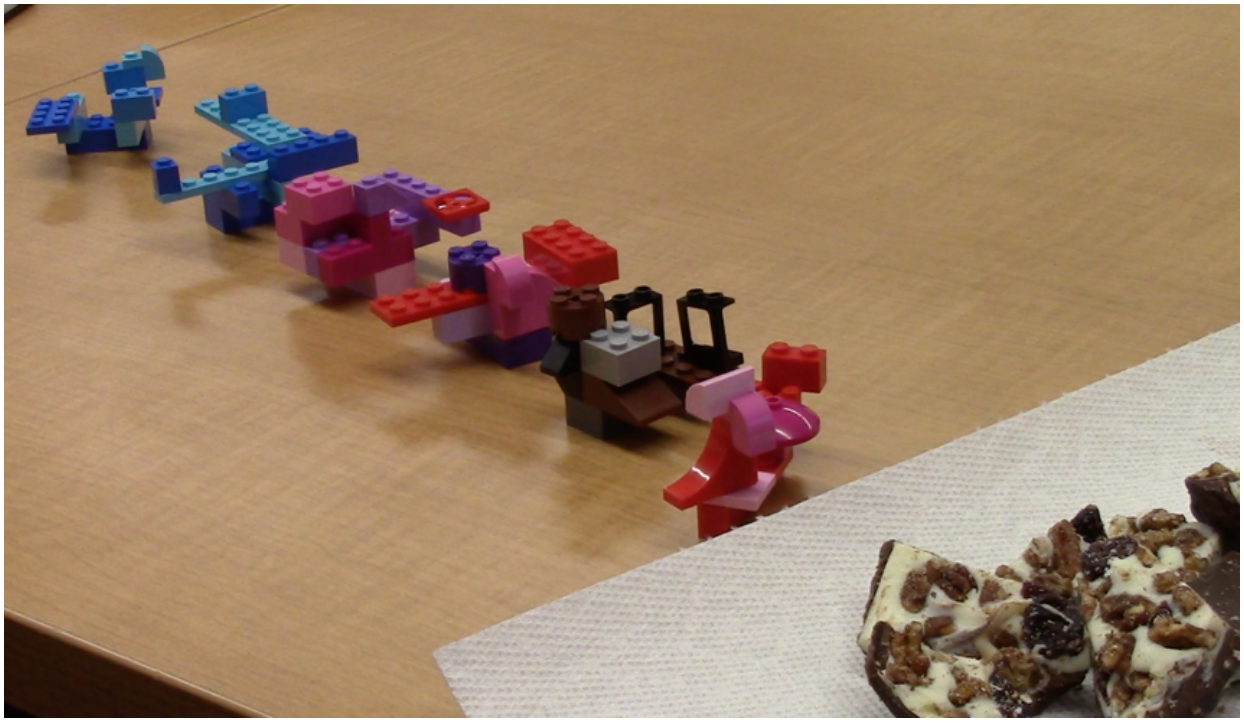
The survey that students had created in October was disseminated in the first week of December and remained open for three weeks. The students received nearly 600 responses, which they analyzed and created data stories from in Spring 2023. To accommodate students' schedules, several small and whole group meetings were convened in January through April 2023 through Zoom. Each of the three-person groups were also asked to attend office hours to discuss their developing data stories in greater depth. In addition, the students were asked to create research journal entries that discussed their research process. I asked them to do so in the form of a story, with a particular emphasis on what worked, what did not work, who were the actors and artifacts in their research, what were those actors and artifacts doing, how were those actors and artifacts cooperating or resisting their efforts, and what possibilities and responsibilities were emerging in their research. These questions were intended to emphasize certain themes in the course so far: that researchers cannot be separated from the phenomena

they research; that the more-than-human world is always an active participant of research; and that research is intended to open up new possibilities rather than foreclose them.

Vignette 1: Legos

It was the last day of the course for the Fall semester. The sunlight that usually streamed into the small conference room had finally left, leaving behind a soft yellow light that reflected off the rectangular table around which we were all gathered. Assorted colors of Lego blocks scattered themselves around laptops, water bottles, and food wrappers. Tucked away in one corner of the table, Saba had put her finishing touches on a series of Lego-built figures set beside one another, along with a chocolate covered cookie that sat next to them (Figure 2).

Figure 2. Saba's Lego-built Data Story



The Lego-built figures, along with the rest of the ongoing activity that evening, was part of the larger sequence of assignments and activities related to the data postcards connected to the Dear Data project by Lupi and Posavec (2016). Saba's Lego-built structure was based on her

week keeping track of how much water she was drinking. Her creation consisted of six adjacent structures (Saturday through Thursday), each of which was made up of two layers. The color of the bottom layer signified the general temperature that day, and the layer above it signified the amount of water she drank and the reason or surrounding context for her water consumption. Small additions to each structure told a further story. In the sixth figure, Saba added a slide because "it just made me think of how my whole day was going really badly in the morning, but then I guess it got better at the end." Pointing to the second structure, Saba explained that its particular structure appeared that way because "I had a lot of stuff to do that day. I was always, I was always constantly moving around. And I believe that's the week that, that was the week of Thanksgiving and everything. So we were prepping for everything that week." The cookie, as it turns out, was also part of the Lego-built structure. Saba explained, "And this is, this one's made out of the candy because that's the day after I took my test on the Friday and I felt really, really good about my test and I was just sleeping in."

It had been a long time since Saba built anything using Legos. She used to play with them with her brother, who is five years younger than her. "We would always play with them and he would always have those cool Transformer ones and Lego Ninjago ones." Because of these past experiences with Legos, the activity brought on feelings of nostalgia for her. "I remember we would always try to build whatever the thing was on the box without looking at the instructions first, and then we'd be like, wait, we have like five pieces missing. Where do these go? That was fun."

I sat there listening to her story while pointing my camcorder at her Lego-built structure. The idea of using Legos connects to an emerging body of work on data physicalizations, which is "the practice of mapping data to physical form" (Bae et al., 2022, p. 1) in order to introduce

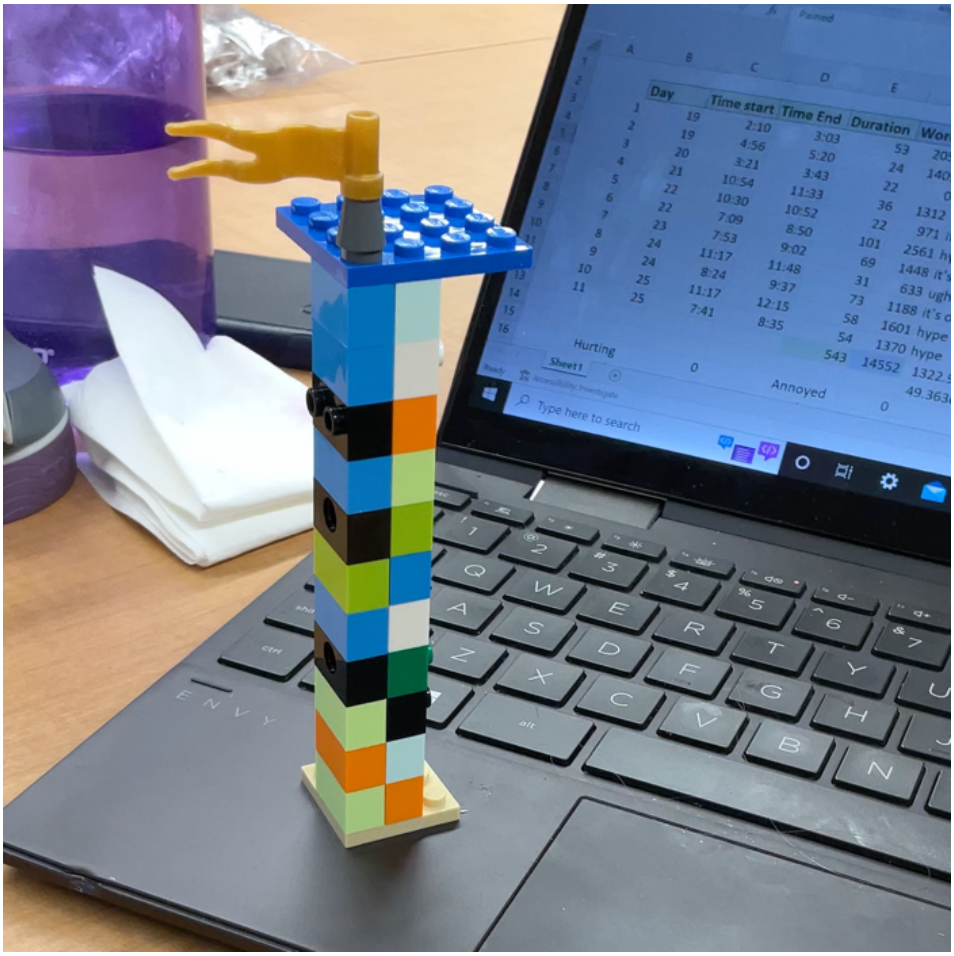
tactile manipulation, accessibility, and immersion in data representations. I asked her if the process of making the data physicalization made her feel differently about herself as a person. In response, she stated:

> I already knew that I wouldn't, or I tend to not take care of myself when I'm stressed out. It kind of made me feel, not gonna lie, when I was doing this like postcard, it made me feel even worse about it because I was like, I'm recording this data and I have to explain that this is the reason why I drank less water that day or I drank more water the next day or whatever.

Earlier, she had told me and her classmates that she had choses to keep track of her water consumption because she knew that whenever she felt stressed, she tended not to drink water or take care of herself as well as when she was not stressed. She was a first year student and had declared major in Biology with the intention to eventually study medicine. Indeed, her response to my question about how making the data physicalization made her feel confirmed her initial hypothesis about the relationship between her stress levels and water consumption habits.

Meanwhile, across the table, Bobble put a flag on top of a Lego-built structure that resembled a skyscraper (Figure 3). The structure was a data physicalization of her fiction writing progress, a topic that she had chosen because she considers herself to be a writer with a particular interest in horror fiction. As part of the data postcard assignment, she kept track of how long her word processing program was open, her physical and emotional state while writing, and her evaluation of her own work. Bobble wanted to make a tower that “wasn’t massive” because the book that she was writing “didn’t feel massive.” But she also made sure that her data physicalization was “not like really thin and small because at the same time it [her novel] doesn't feel really thin and small.”

Figure 3. Bobble's Lego-built Data Story



Bobble talked about her initial feelings about the data physicalization activity, noting that it was "frustrating at first because I'm more of a words guy and trying to, in my brain it's like trying to cram words into blocks, which doesn't, you know, it didn't work great at first until I realized, hey, these aren't words. You've just got to make 'em blocks and work with the blocks." In the previous iterations of her data postcards, she felt compelled to make sure that her visualization formed a one-to-one correspondence with the event that she was measuring. In her postcard, "there is a bar representing like 80% of it is full, so that means I did 80% of the work." The blocks, however, were more freeing because, as she stated, "these are just blocks. It doesn't, I don't have to make something that looks what it is supposed to mean." Indeed, the blocks "kind

of detached me a little bit from like all the fiddly little statistics and Excel sheets and went, like, turned into more of like, what it was, like a block of time that I had an experience in instead of just a bunch of numbers and words that I ascribed to them."

Moments later, Bobble's hands began to move across the structure and broke it apart (Figure 4). One Lego-built structure fell into a handful of pieces, some scattering haphazardly beside the laptop and others falling dangerously close to the edge of the table. She began to talk about her writing process once again: "At the end of every book I decide I hate it and I never return to it ever again." Describing her action as a metaphor, she explained: "I'm destroying all of, well not destroying, but I am taking apart all of the work I did and now putting it into a thing that no longer has the meaning I ascribed to it."

Figure 4. Bobble Disassembling Her Lego-Built Data Story

I asked her how it felt. “Cathartic. A little bit cathartic. Yeah.”

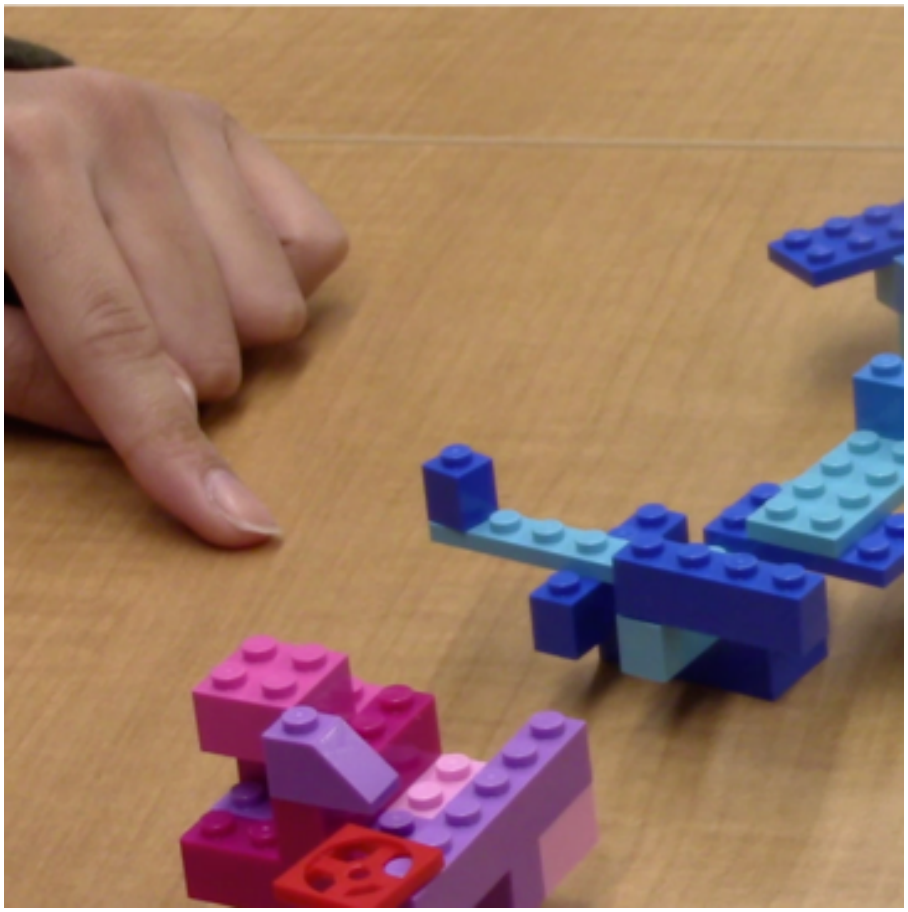
Sub-Section 1A: Data, Bodies, and Representationalism

It is common in a course or lesson about data to discuss the topics of data *collection* and data *visualization* (and more broadly, data *representation*). These terms—collection, visualization, representation—can suggest that data is an abstract concept capable of being separated and pulled away from the phenomena that data purport to describe. In this way, data is a *pointer* to a reality that we might otherwise not be able to access. In computer science, a pointer is an element within a programming language that stores the memory address of another variable. For instance, the variable *x* might be assigned a numerical value, such as 5. If the pointer *a* is assigned to *x*, then *a* can be used to call upon the value of 5. However, manipulating *a* typically does not change the value of *x*. If the value of *x* is changed independently, then *a* continues to point to *x* but now calls upon the new value assigned to *x*. Likewise, the pointer *a* can be reassigned to a new variable *y* without impacting *x*. This allows the pointer *a* to exist independently of *x* and to be used for computational purposes without affecting the world that it references.

This orientation toward data might be regarded as a kind of neo-platonism, which expresses the “belief in a hidden mathematical order that is ontologically superior to the one available to our everyday senses” (McQuillan, 2018, p. 2). This hidden mathematical order consists of parameter values stored within the variable *x*, and data becomes the pointer *a* that enables us to access the contents of *x*. D’Ignazio and Klein (2020), citing Haraway (1988), would describe this as the “god trick” of data. The god trick creates the impression that people can use data to understand the world from a detached, distant, disinterested, and disembodied viewpoint, much like a god who sits above and away from the everyday world.

Throughout their process of creating their data postcards and Lego-built data physicalizations, Saba and Bobble (along with their peers) regularly expressed the ideas of data *collection* and data *representation*. These discourses of collection and representation were entangled with the materiality of their Lego-built structures and data postcards. The ideas of "collect" and "represent" were not words alone but rather discursive-materials that participated in boundary drawings that created the conditions for data to emerge as an abstract pointer that referenced a separate and independently existing reality. As Saba pointed to each segment of her Lego-built structure (Figure 5), she enacted several separations. First, as she was speaking to me about her Lego data physicalization, she used her finger to gesture to and her hand to grasp individual sub-structures, sometimes sweeping her hand from left to right or holding out her hands to present the sub-structures as a collective whole. These gestures, along with the configuration of the Lego data physicalization itself, which was arranged from left to right in front of her body, enacted a separation of time into discrete days that together formed a single week. Each of the Lego sub-structures thus became a stand in for her water consumption activity on each particular day. As she made statements such as "the blues represent cold temperature. The reds and pinks represent the warmer temperature outside," she separated the concept of temperature into discrete categories. Each of these categories mapped to a different color, suggesting that the phenomena of temperature could be "captured" through color and subsequently "represented" with Legos. The Legos themselves were not data but rather became a physical substitute for data. Data were instead the immaterial idea of days and temperatures. Data-as-idea are capable of being translated into the material world, but nevertheless they retain their form as an abstract pointer referencing an independently existing reality.

Figure 5. Saba Gesturing Toward Her Lego-Built Data Story



Saba's words and gestures, along with the color-coded and spatial configuration of the Legos, created the conditions for Saba to make causal claims about the world. While pointing to and talking about each of the colors, she stated that the temperatures on each of the days "affected how much water I drank those days. There's a lot more stuff like red over here because that's the day that I had looked to study for a test the next day. And then this one is, um, it has a lot more black on that day because I drank a lot of water that day because I think it was like the perfect temperature outside." The gestures, physical configuration of the Legos, and assignment of color and space to temperature and days all created the conditions for her to take a step back and make an "objective" statement about her water consumption patterns and the reasons for them. This assemblage of data momentarily positioned Saba as a participant in the god trick of

data. Importantly, this positioning did not occur through discourse alone. It was not the case that Saba merely expressed or was shaped by broader discourses and storylines about data, where such discourses and storylines can be understood primarily in terms of language. Instead, as this analysis suggests, the idea of data-as-pointer emerged through discursive *and* material configurations within a data assemblage. There is an irony, which is elaborated further, that data's *materiality* can make possible the perception of its *immateriality*.

Saba and her Lego-built data physicalization were important forces within the classroom, but they were not alone. In a later interview, Bobble stated that in relation to her data physicalization, "I consider data a concept…Data doesn't really exist. It's sort of our interpretation of something in reality." Like Saba, Bobble pointed to separate pieces of her Lego-built data physicalization while explaining what each part "represented." She accompanied each gesture with the explanation that "each of these represents one data point I took…So the blue was happy, you know, this gray or this, this, this color is annoyed. And then I have green for apathetic. The orange was sated and then lighter blue for ecstatic. Oh, the black is annoyed. But um, so I guess I changed it." I do not look at words alone but rather the discourse-materials within a data assemblage in order to understand what boundaries are being formed and therefore what possibilities are being enacted. In this case, Bobble quantizes the emotional responses that she felt after writing her novel into discrete categories. Color becomes the mechanism by which these categories are expressed. Just like color enabled Saba to speak to her patterns of water consumption "at a distance," the colors of the Lego blocks that Bobble gestures toward enable her to speak about her own emotions with a degree of "objective" detachment. The Legos are a mechanism of communication and a physical manifestation of her interpretation of an independently existing reality. Data is once again made to "not really exist." It is a mere pointer

to something that does exist, and that pointer manifests in different forms according to one's interpretation of an event.

The circulation of bodies and discourse-materials that included Saba, Bobble, and their Lego-built data physicalizations were part of a data assemblage that also included the course on data storytelling itself. When considering how the circulation of bodies and discourse-materials within a classroom disrupt or interrupt onto-epistemological assumptions about data, one cannot separate the students from the broader classroom environment that includes the teacher(s) and their pedagogical and instructional design practices. Barad (2007) emphasizes that there are deep onto-ethico-epistemological entanglements between the "objects" of observation and the "agencies" of observation. Both are co-constitutive within phenomena. My purpose is not to be *reflexive* in my role as the teacher and researcher in the data storytelling course but rather to attend to my response-abilities with respect to the boundary making practices that occurred in the course. In the initial assigned directions for the data postcard activity, I gave the students the following instruction:

> In this exercise, you will be asked to collect data about yourself over a period of time and reflect on the process of data collection, analysis, and representation. The purpose of the activity is to learn about the data collection and representation process, to reflect on the four elements of data (artifact, processable information, interpretable text, and evidence), to consider how power may or may not come into play in data collection and representation, to understand the role that data can play in your everyday life, to think about story structure, and to experience a low-tech approach to data storytelling.

In a discourse-focused analysis, one might interpret these directions in terms of positioning Saba and Bobble to act in the role of data collectors and representers, thereby pushing them to adopt the god trick of data. However, this analysis, which thinks with socio-material scholarship in order to conceive of data as an assemblage, seeks to go beyond the discursive. The students and I, along with all of the other bodies and discourse-materials circulating in the classroom, were

co-constitutive of the enactments of data that occurred in the course. Despite my sustained interest in scholarship that tends to eschew the notions of data "collection" and "representation," I generated the above text not because of a "mistake" but instead as a part of the onto-epistemological legacies of data that endure within the discourse-materials of data assemblages within which I acted as a teacher and researcher. The question about which this analysis is concerned is not to focus on particular causal explanations for why the students and I took the particular actions that we did in the course. Rather, it is to become oriented to how the circulation of bodies and discourse-materials produce possibilities, and then to consider what response-abilities we bear on account of these enactments. The next section is my effort to take up such response-ability by asking what else might have been possible in the course beyond reproducing the onto-epistemological assumptions about data that have been discussed so far.

Sub-Section 1B: Data, Bodies, and Non-Representationalism

Data are local phenomena constituted not only by the words and actions of people but by the agentic capacities of materials. As Loukissas (2019) states, "data are cultural artifacts created by people, and their dutiful machines, at a time, in a place, and with the instruments at hand for audiences that are conditioned to receive them" (pp. 1-2). While Loukissas (2019) adopts a humanist approach around the agency of "dutiful machines," Ingold (2011) goes further to say that tools and other materials are equal agentic partners that act alongside, rather than dutifully subservient to, the hands that wield them. He states, "materials…are the active constituents of a world-in-formation" (2011, p. 28). Referring to an example about the agency of a stone, he explains that "stoniness" is not "merely in the mind of the observer or practitioner. Rather, it emerges through the stone's involvement in its total surroundings – including you, the observer – and from the manifold ways in which it is engaged in the currents of the lifeworld" (Ingold,

2011, p. 32). In other words, "human beings do not exist on the 'other side' of materiality, but swim in an ocean of materials" (Ingold, 2011, p. 16).

In the previous section, I discussed how the arrangement of Saba's Legos, along with her gestures and descriptions, enacted an onto-epistemological separation between data and the world that data point to. Here, I consider the possibility of reading the same moment as Saba "swimming in an ocean of materials." In this reading, I attune to the materiality of data in its involvement with the bodies and discourse-materials circulating in Saba's Lego-built data physicalization and the impact that these bodies and discourse-materials had on each other. As Saba was talking about her Lego sub-structures and discussing her thoughts about how it felt to create them, a strong affective response began to emerge. She held out her hands in a gesture of presentation, explaining that she felt bad about the activity because it required her to confront her reasons for drinking less water—namely, the fact that she drank less water when she was stressed from school. She did not necessarily carry these feelings with her into the classroom as if they existed all along. Instead, her data physicalization, the table, my presence, and the quietude of the room seemed to have produced this in/with/through her. Each Lego sub-structure, though seemingly inert on the surface, told a different story, richly textured and defined by Saba's personal feelings and her awareness of the connection between her health, her professional aspirations, and her life as a student.

The linear arrangement of the data physicalization created the possibility for a surprising element to emerge—the cookie that lay at the edge of the data physicalization. Bobble and her family had made the cookies at home and brought them to the class to celebrate the end of the semester. On the same day, my co-instructor and I had brought in food from a local restaurant based on students' requests that they had made ahead of time. By the time students were creating

their data physicalizations, the single conference table on which everyone worked was strewn with napkins, plastic bags full of cookies, half-torn wrappers, laptops, cell phones, loose Lego pieces, and Lego-built data physicalization structures that were in process of being built. This seemingly haphazard configuration of bodies and discourse-materials created the possibility for Saba to engage in a practice of improvisation, whereby she incorporated a found object into her data physicalization. The cookie was not a "dutiful" object ripe for manipulation but rather an "active constituent" in a "world-in-formation," alongside the Legos that lay next to it. Noticeably larger than the Lego sub-structures, the cookie was Saba's celebration for having positive feelings after taking a test. The cookie was not readily manipulatable, did not have colored layers to signify days or temperatures, and its overall function as a data *representation* was unclear. It became a spur of the moment addition to the data assemblage, subverting the supposed permanence of data on account of its status as a perishable good. Moreover, it was not the case that Saba merely "used" the cookie in her data physicalization. As Ingold (2011) states, the use of an object does not arise from the act of attaching an attribute to that object. Instead, use entails:

> [J]oining a story to the appropriate gestures. The tool, as the epitome of the story, selects from the compendium of the hand the gestures proper to its re-enactment. Yet the tool has its story only because it is set in a context that includes the trestle, the wood, and all the other paraphernalia of the workshop. And the hand has its gestures only because it has grown and developed within the organic synergy of practitioner, tool and material. (Ingold, 2011, p. 58)

The cookie had its own story, and on account of this story, had agentic capacity in its intra-action with Saba. Cookies are widely regarded as delicious and often associated with feelings of happiness or reward. This is because cookies are usually meant to be eaten. In Saba's case, the attributes of the cookie conformed to the situation in a new way, taking on the role of a participant in a data physicalization about water consumption, academic performance, and stress.

Thus, the trait of being "delicious" was not an inherent attribute of the cookie but rather a story that co-emerged with and acted on Saba, just as much as Saba acted on it by picking it up and placing it next to the Lego sub-structures. Thus, Saba, the Legos, and the cookie had "grown and developed within the organic synergy of practitioner, tool, and material." This highlights a degree of intimacy with data that is distinct from the typical view that people *use* tools to *collect* and *represent* data in order to say something about the world from afar.

McQuillan (2018) invites readers to consider an agential realist approach to data science that seeks to understand how data might be considered "an apparatus whose role in 'sedimenting reality' is open to participatory reworking" (p. 16). I turn to Bobble and her data physicalization to consider this possibility of participatory reworking. Following my initial conversation with Bobble, I returned to her work area to find that she had begun to disassemble her Lego structure. Rather than this moment signifying the conclusion of her data physicalization, it signified the beginning of a new story. Before she disassembled the structure, she had noted that she did not make the tower "massive" because the fiction story that her data postcard was about did not feel massive. Likewise, she did not make the structure "thin and small" because her story did not feel thin and small. As she began to disassemble the tower, she related this act to the emotions she felt each time she finished a piece of fiction writing:

> We can relate this to the book cycle. That thing I was about in my second data story thing where, uh, at the end of every book I decide I hate it and I never return to it ever again…I am taking apart all of the work I did and now putting it into a thing that no longer has the meaning I ascribed to it.

She then described the experience as "cathartic." Bobble and the Legos enacted a participatory reworking of data. Data were no longer an abstract pointer to an independently existing realty of quantifiable emotions. Instead, Bobble's acts of creation and destruction were themselves data. Such data were not pre-planned. They were not the product of deliberate action taken in order to

"represent" the world from a distance. Instead, such data were the result of Bobble's hands shaping the Legos just as much as the Legos were shaping Bobble's hands. From this shaping emerged Bobble's feelings about her books, including her feelings of hatred toward any book she finishes writing.

The data-as-Legos, in other words, had a performative effect that often goes unacknowledged when data is seen as an abstract pointer to a separate reality. Data induces a response on the world rather than merely measuring it. In Bobble's case, this occurred because of the *story* that her acts of creation and destruction told. Stories are the mechanism of agency by the material world. "Things are their stories" (Ingold, 2011, p. 75). Oppermann (2021) explains that because "matter is agentic and capable of expressing itself, it must have a narrative dimension; it must be *storied matter*, a living text encoded with meaningful signs and/or creative expressions" (2021, p. 265) . When I asked Bobble what she thought was meant by the idea of "story," she gave the following responses:

> **Lee Melvin**: In your view, what is a story?
>
> **Bobble**: I'm gonna give you the straight up…anything. This chord, it's got a story. It can be a story. I can make this into a story. Uh, this list, it's a story. It was from a couple of months ago. This was me listing the things I wanted to do, um, for my birthday. That's a story. That's definitely a story. Uh, that lamp that appeared in my room while I was gone, that's a story…Um, the data of physicalization is a story…anything and everything. It is a story, whether you like it or not. It's multiple stories.
>
> **Lee Melvin**: So, is there anything that's not a story?
>
> **Bobble**: No. No. I mean, you could take any object in the world and it's got a story. Like, even if it just sat in one place for its entire life, that is a story, and that's poetic. Sorry, but yeah.
>
> **Lee Melvin**: What gives it, what makes something a story?

> **Bobble**: Uh, I guess sort of existing. Cause, I mean, I consider a story just the tale of something's life, sort of. So, I mean, by existence? Yes, being there. I don't really have a solid answer to that one, but it's there, so it's got a story.

In this dialogue, Bobble offers an ontological account of stories. This account resonates with those of Ingold (2011) and Opperman (2021) in suggesting that all matter *are* stories simply on account of being. The Lego sub-structures, the cookie, the Lego tower, and the crumbling Legos were all storied matter. Through these stories, matter enacted their agentic capacities. Data not only *expressed* ideas about the world but more importantly *acted* on/with/through/in it, shaping feelings about personal health, academic performance, stress, personal accomplishment, attachment, and resentment. Data took on the stuff of matter by becoming part of the gestures, hands, and bodies of the students, the instructors, and the classroom. These bodies and discourse-materials became part of larger data assemblages, some of which reproduced and others of which disrupted onto-epistemological assumptions about the materiality of data and its relation to the world.

Like Bobble, whose act of breaking apart her Legos began a new story in place of a conclusion, I close this section by sharing the original data postcards that Saba and Bobble created, as well as stories that I had asked them to write to think speculatively about the agentic capacities of the materials involved in the creation of those data postcards.

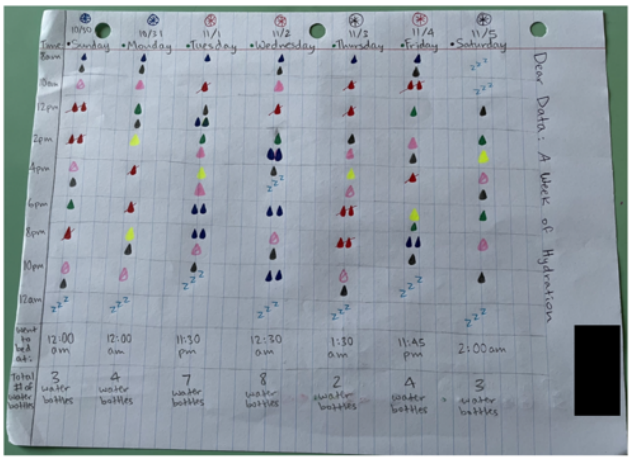
Figure 6. Saba and Bobble's Data Postcards and Data Postcard Stories

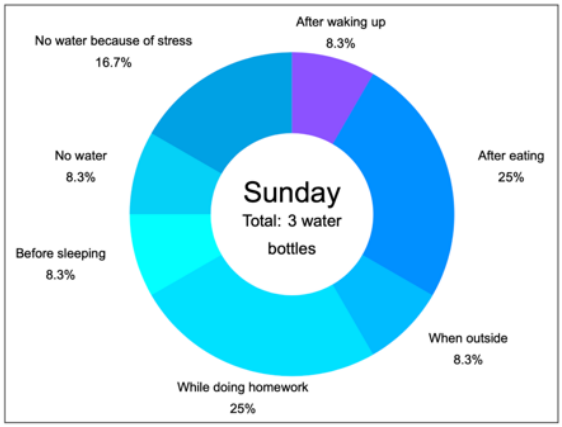
In *Alien Phenomenology, or What It's Like to Be a Thing*, Bogost (2012) invites readers to consider the phenomenological implications of an object oriented ontology. He asks what it might mean to consider a non-human world that exists outside of our own conceptions. Is it possible for humans to "get out of our own way" and consider the experiences of hammers, nails, and microcomputers outside of their use for us. That is, "what is it like to be a thing…on its own terms" (Bogost, 2012, p. 10). Inspired by this work—which seeks to challenge what Meillassoux refers to as correlationalism—I asked the students to write brief stories about one or more non-human materialities that played a role in the creation of their data postcards. In other words, based on a "philosophy *claiming that things speculate*," I asked them to "*speculate about how things speculate*" (Bogost, 2012, p. 31, emphasis in original).

Figure 6 (cont'd)

Below are the alien phenomenological stories written by Saba and Bobble. Each of their stories are accompanied by a sample of the hand-written and digitally-produced data postcards associated with the materialities about which the stories were written.

Saba



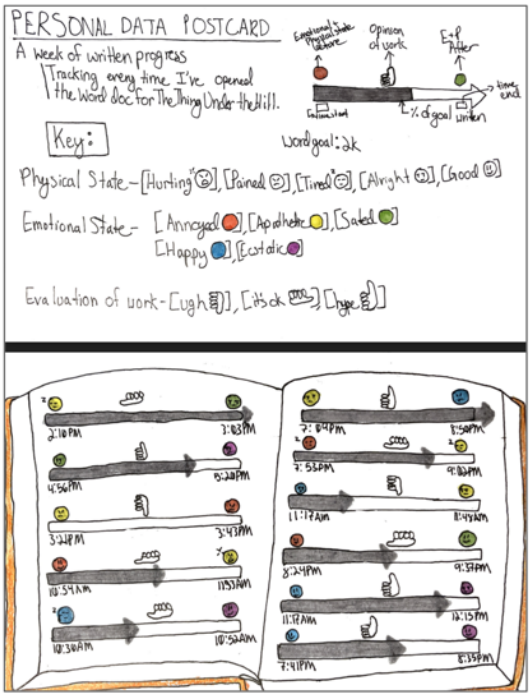


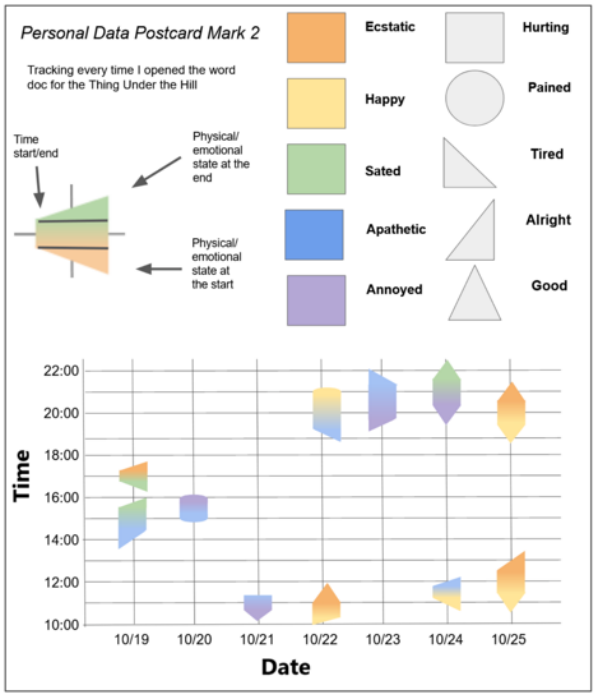
Finally! We are being removed from the box that has held us for so long. After a while, we almost thought we would never leave. One of us even thought we would dry out soon. Click! What's this? Blue is being used a lot and I think Purple is very jealous right now. Click! Click! Now Red is being used to cross out these water drops? The rest of us wonder what for.

We're all slowly being used to draw more and more water drops. Red was very tired by the end, but the rest of us were very proud of Red. At the end, we all saw that the different colors were being used to represent the reasons why she drank water at different times throughout the week.

Figure 6 (cont'd)

Bobble





I am used to wipe away the wrongs.

I am carefully brought down onto the page, held at an angle, to cover up the areas she thinks do her a disservice. The numbers that go backwards, or the lines misplaced on the page. I drag over misspelled emotions. The remains are white, too bright for the papers true color, shiny and easy to smudge when she attempts to replace 'anaesthetic' with 'apathetic'.

It's odd, to me, the things she chooses to hide. Nobody will see this paper I am on. She hesitates when she writes the time, always, glancing at the clock and wondering '*did those minutes really pass? did I really spend that long staring at all that nothing?*'. I sit at her elbow and wait.

The day's slides watch both of us from the computer screen. I think I see her look at them, out of the corner of her eye, and I bet she wishes she never closed that Word Doc now. I am held. Unscrewed, carefully.

Maybe she'll add more words to her count. An extra five hundred mean nothing to the faces who will see it– who could know? Who could tell, in this page that I dot like a plague? Her hand hovers over the numbers. I can feel the tremble. The decision.

But she does not hide them away. Nobody would know, and nobody would understand, but the ducks would. The ducks would see it, the ducks would understand, and although they would

Figure 6 (cont'd)

not judge (The ducks have seen worse. The ducks have seen it all), she doesn't want to disappoint them.

Or disappoint herself. So I do not cover that 971. It stays that way, even when she sees it as a percentage, and curses at her hands.

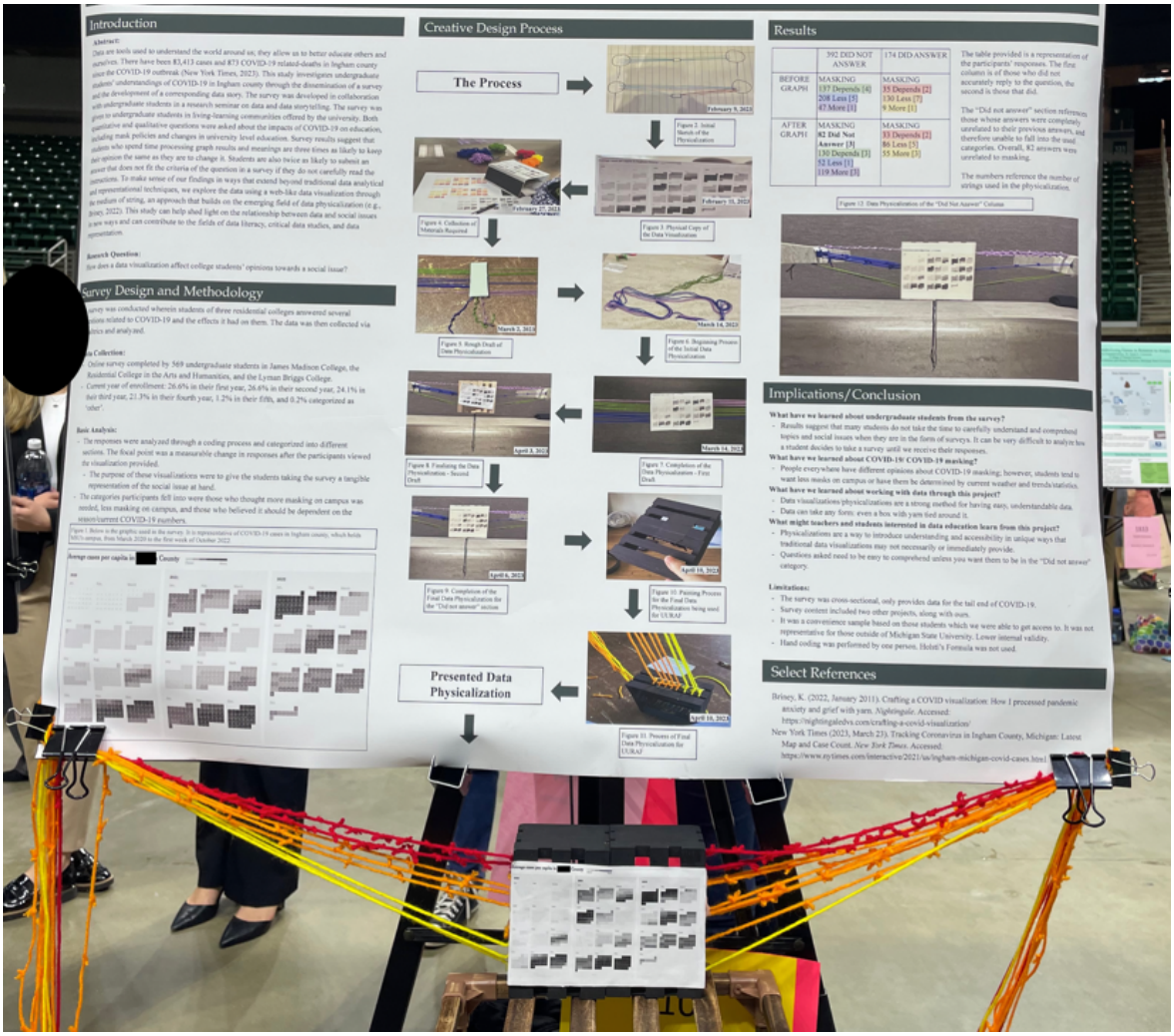
She only ever wrote those 971 words, and I am used to wipe away the wrongs.

Vignette 2: Yarn

Bobble, Saba, and Daisy just arrived at the arena floor where an undergraduate research conference was taking place. They were scrambling to set up the data physicalization that they had created as part of their year-long research project for the data storytelling course. The data physicalization was made out of different colored strands of yarn interrupted with knots throughout and held together—or better yet, holding together—a wooden block on which was affixed a printed black and white image displaying the average number of cases of COVID-19 in their county (Figure 7).

The yarn, much to their dismay, kept slipping off its intended support structure. Bobble, Saba, and Daisy scrambled to figure out a solution. They improvised, darting from empty poster board to empty poster board around the arena and gathering enough clips to anchor the ends of the yarn to the legs of the easel on which they mounted their own research poster. In the end, they sighed a breath of relief as they finished setting up their station right in time for the research conference judge to come around their poster and start asking questions.

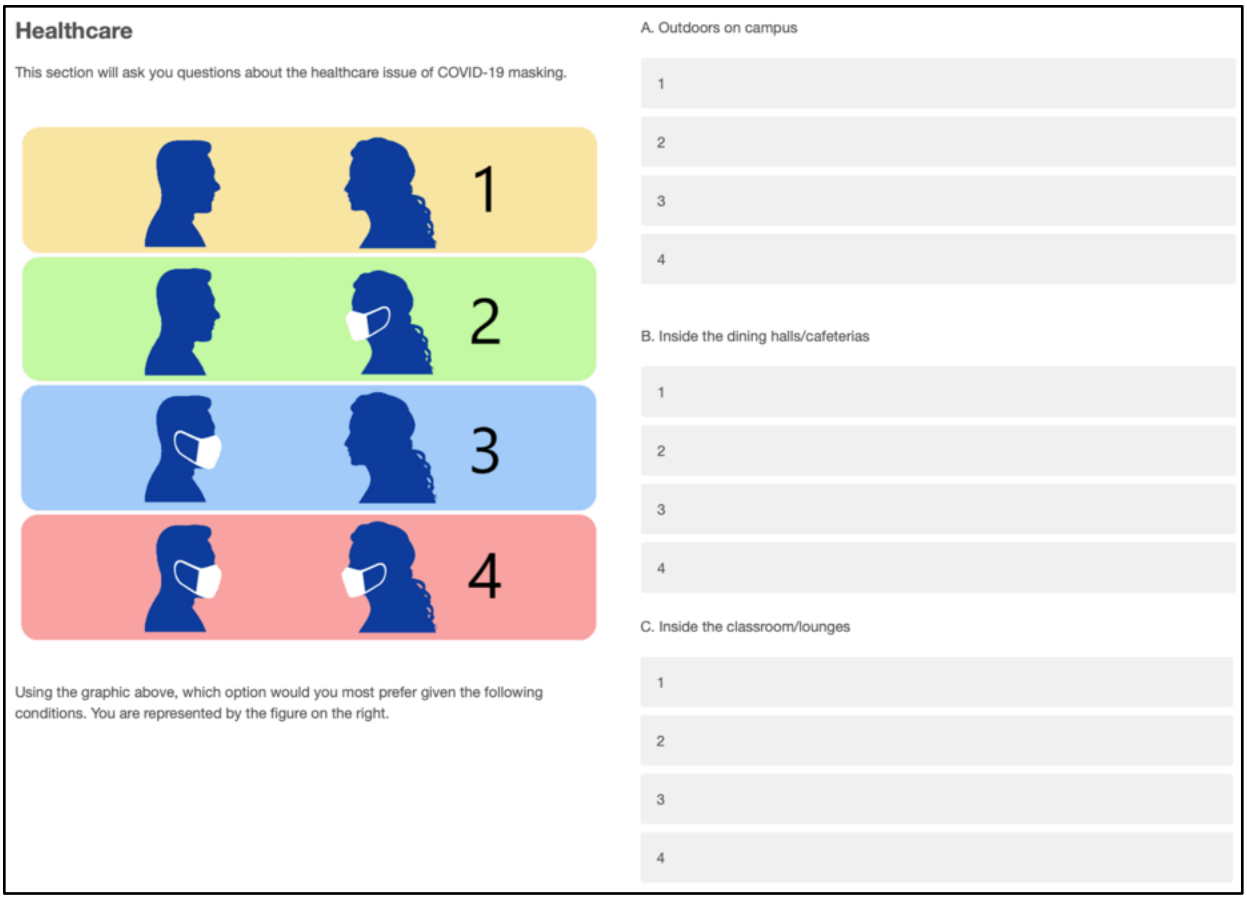
Figure 7. Bobble, Saba, and Daisy's Yarn-Based Data Physicalization and Research Poster

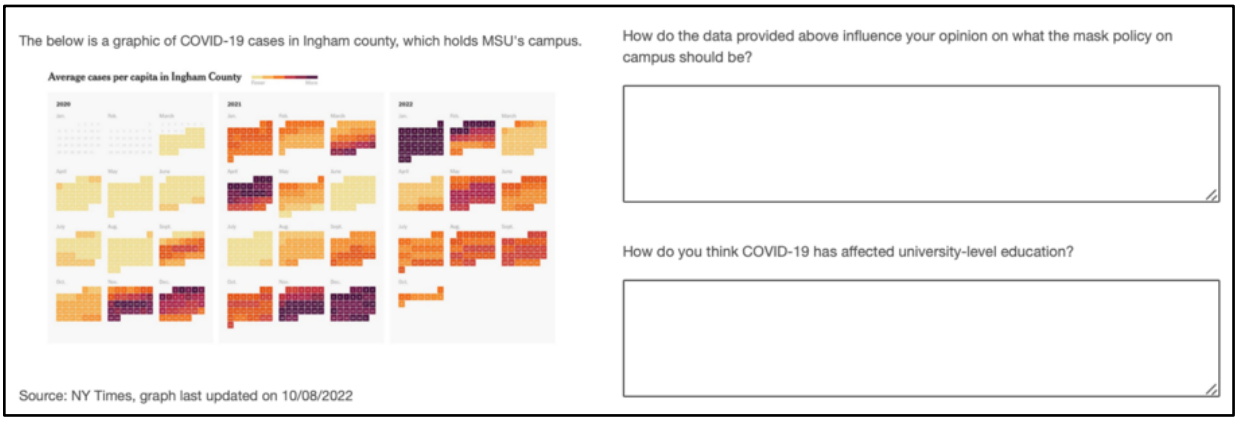


Earlier in the previous Fall semester, Bobble, Saba, and Daisy had chosen the topic of COVID-19 masking as the topic for their research project. The parameters of the assignment were as follows: each group of three was asked to choose a social affairs issue, to design survey questions around that topic, to disseminate their survey to students at the university, to analyze the results, to create a data-driven "story," broadly construed, based on their data, and to present their research at the undergraduate research conference that would take place the following Spring. They chose the topic of COVID-19 because Saba and Daisy wanted to attend medical school and thus had an interest in healthcare issues. Bobble had a personal interest in COVID-19

as well. They found and modified an infographic about COVID-19 masking in order to design a survey that asked questions about students' masking preferences in different situations. They also borrowed a graphic of COVID-19 cases in their county to understand how the data visualization impacted students' opinions about masking (Figure 8).

Figure 8. Bobble, Saba, and Daisy's Survey Questions





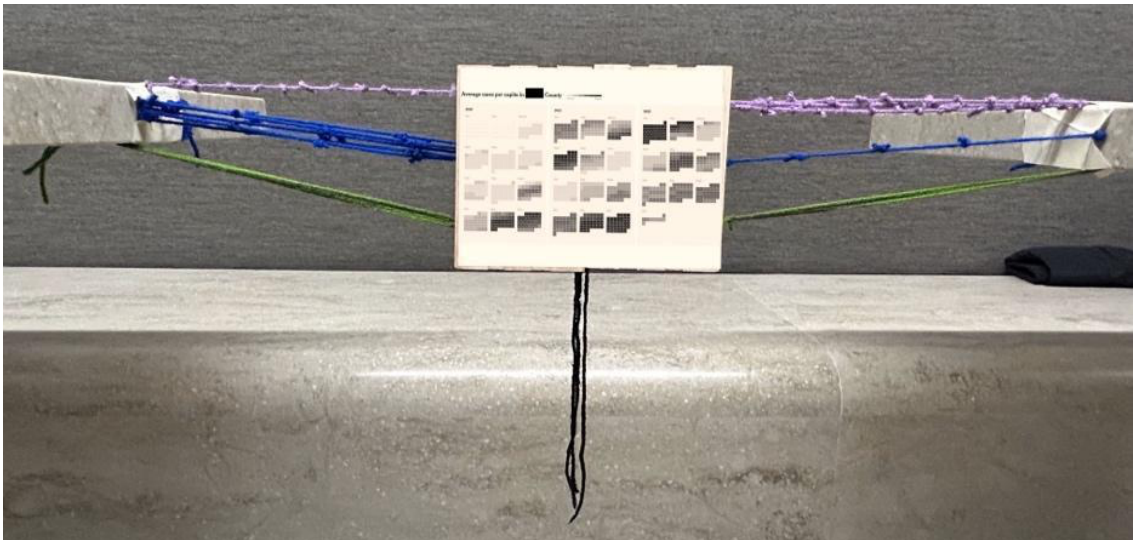
The group had decided to make their data story in the form of a data physicalization. As Bobble explained in December:

> I have a vision in my head of a bunch of strings, sort of, in a webbish sort of thing, color coded (because I love color coding) and spun in a certain way that shows the data. I have absolutely no idea, but I'm invested in the idea of making our story look like COVID– the web of the data … I also really like the idea of this being an experience in and of itself: like the people walking by can put stickers where they lay in the visualization, like we're collecting data as we present our data and layer it like a good croissant.

For several months, the group created prototypes of potential design (Figure 9). If their initial idea did not work, they threw it away and started another one. Each idea used yarn in some way, with the final iteration using colors and knots to stand in for people's opinions about COVID-19 masking. For instance, strands that had many knots stood for respondents who answered mostly "it depends" as to whether they wore a mask or asked others to wear a mask as a precaution against COVID-19. The group explained that the knots were created for accessibility purposes to enable people who are visually impaired to interact with the data physicalization.

Figure 9. Prototypes of the Yarn-Based Data Physicalization

Figure 9 (cont'd)



Referring to the day of undergraduate research conference, Bobble recalls how she felt when trying to set their data physicalization up:

> I was very stressed out, and I think I remember you guys telling me to stop stressing out over everything as we set up the poster, and we didn't have any clips on our board to put up our poster, and then, as we were trying to tie up the physicalization, it kept falling down, and everybody else's was going up just fine, even because they had less, less parts, less finicky pieces, and the yarn kept slipping because it was, it was yarn and not like string, so no matter how tight we tied it, we couldn't get it to stay as a knot, so we ended up taping it down, and, but it got upright. And then I was stressing over that because I wanted it to look so polished because I was afraid people were going to judge it for being messy.

However, when the research conference judge came around, he did not seem to notice.

> They were so interested in a physicalization. It was like what I, like, I thought it was messy, because I, I saw the effort that I put into it. I saw how the yarn kept slipping. I saw all of that, but they didn't see that. They just saw a physicalization, which was something that I don't think any of the other pieces had. It was the physical representation of data and they were so fascinated by the idea of having a physicalization that they didn't care that it was that, like, they didn't see the issues that I saw in it … He didn't know that the yarn slipped. He didn't know that it was supposed to be suspended. He just saw that we had an interesting idea, and we executed it. And he wasn't judging me for it. He was asking questions about it because he was genuinely interested in what we'd done. And it was a really nice experience of realizing, yeah, I had done a lot of work for this.

Indeed, Bobble likened the entire conference experience to a data physicalization itself:

> I didn't really believe it until I saw a physical version of that qualitative data of the judge being interested in it, like, in a way that I could, like, interact with, which kind of reminded me of our data physicalization, how there's, it's one thing to have qualitative or quantitative data that you can look at, but it's an entirely other thing to be able to interact with it, if that makes sense.

Later on, I asked Bobble if creating the data physicalization changed the nature of the data itself. She responded that it did "because you got all sorts of bias leaking from all of your pores. And, like, I had to code it. That's, in order to physically represent the data, I had to put it into categories. And that's completely changing the thing, because Saba could have done it and she could have had completely different categories than me."

Bobble then started talking about the relationship between their research project, data physicalization, and the issue of bias in data:

> You're putting your bias all over it and you're putting it into categories, which is automatically challenging the integrity of it, but it's also necessary in order to create a representation … And it's never going to be absolute, like sure you can, you can interview everybody at [redacted] and still would not get a truthful interpretation because you oozed your bias all over the questions and all over how you collected it and you did it through a computer or through a tablet or anything like that and then after you've got all that oozed over collection, and then you try and put it in categories, which data resists because it's automatically not a thing that exists. So you're putting a concept into a bunch of other category concepts. And then after that, you're trying to turn it into a physical entity, and in the creation of that, through the materials you use, through the people you use it, through the time you do it, you are also oozing bias in your own interpretation onto it. So you've got an interpretation of interpretations of an interpretation of an interpretation of data. So it's barely the same thing.

Ultimately, Bobble suggested that the data physicalization was a way to tell their data-driven story in a more humanizing way:

> It was less stuffy, it was more human. I felt like people were, felt more comfortable asking us questions because we obviously had enthusiasm for the subject and we were putting, we were getting our human all over it. Like it wasn't a graph where graphs, like a lot of graphs are made to look impersonal and unbiased and correct and medicinal almost in like the grid squares and shit. Um, but we had a bunch of yarn tied to a box. And that automatically that I think that made people more comfortable with both the data that we were representing and

> like the people presenting us and presenting it and like what we were trying to tell them. I guess character of the box being more human and well worn affected how people would consider it or like the narrative of the box and in turn that affected how people considered the narrative of the poster because it was tied to the yarn box. There's levels there. I don't know.

In fact, Bobble indicated her own connection to the data physicalization:

"I damn well can connect, like, pack bond with a box, and I did, I love that thing."

Sub-Section 2: Data, Time, Risk, and Uncertainty

Data often involves the creation of variables, coding variables, and partitioning variables into equally sized groups or equally spaced intervals. In a scatter plot or bar graph, for instance, the *x*-axis is sometimes assigned the variable of time. In a typical data literacies course, an instructor emphasizes that if there is an *x*-axis representing time, students must ensure that the tick marks drawn along the *x*-axis are equally spaced and arranged in chronological order. In 2020, the Georgia Department of Public Health released a graph about COVID-19 cases where the *x*-axis violated both conventions. The *x*-axis was not arranged in chronological order, let alone partitioned along equally spaced intervals, with the result that COVID-19 cases appeared to be declining in Georgia over time. This graph led to accusations of bias and a corresponding apology from then Georgia Governor Brian Kemp, who blamed a third-party vendor for the "mistake" (Sabin Center for Climate Change Law, n. d.).

There is an assumption—and moreover, an insistence—that whenever a person engages with time-based data, they must partition time into equally spaced sub-divisions and treat time in chronological order. Mikulan and Sinclair (2023) discuss the notion of "Time-as-Arrow," where time is "abstract, one-way, singular, and teleological" (p. 2) and change occurs within a backdrop of time that is "empty, universal, a homogenous vessel" (p. 32). This is the dominance of "clock time" (Duncheon & Tierney, 2013, p. 237), which conceives of our lives as occurring along a

continuum of absolute and discrete moments that can be objectively identified and labeled according to a system of time-keeping that is universally shared across all living and nonliving things. Clock time carries the assumption that "quantifying time allocations provides adequate understanding of an individual's relationship with time" (Duncheon & Tierney, 2013, p. 243, citing Aminzade). Linear and uniform subdivisions of time "might be thought of as a form of 'compulsory hetero temporality'" (Rifkin, 2017, p. 39). In this way, "temporal discourses are discourses of power and resistance" (Mikulan & Sinclair, 2023, p. 37).

Before discussing Saba and Bobble's work with their yarn-based data physicalization, I return to a discussion of the data postcard activity in order to highlight where notions of clock time and time-as-arrow emerged in the course. Saba's data postcard displayed her water consumption along an *x*-*y* grid of days and times. This format is common in data visualizations that express change over time. It creates and imposes the effect of clock time onto data visualizations in order to present data as capable of enabling people to identify patterns, draw conclusions about phenomena, and make predictions about the future. Saba stated, "I never really thought about what data actually was and why it matters. Now I understand that data are simply information that have been translated to be more efficient and understandable." When read alongside the grid-like structure of her data postcard, her words are not merely a reflection of her beliefs about data. Instead, they enacted the possibility of her data postcard to become a tool for "efficient" understanding. Likewise, her data postcard enacted the possibility for her statements about efficiency and understandability. Both the illustrations in her data postcard and her discourse co-constituted one another within a larger data assemblage that included not only the materiality of data (as discussed in the previous section) but also its temporal dimension.

In addition to time-as-arrow, Mikulan and Sinclair (2023) invite us to consider other provocations about time, such as the notion of time as repetition. As part of their discussion, they refer to the "grid-like space-time coordinates of representational conceptions of difference and repetition" (p. 124), the "qualitative repetition of cosmic time" (p. 122), the "rhythms and demands of the capitalist, neoliberal machine" (p. 127), and "efforts at organizing, mining, extracting and managing the rhythms of phallic figures such as 'nature', 'culture' and the current 'techno-oriented' modes of capitalist production" (p. 138). These references refer to popular conceptions of time where 'history is bound to repeat itself'. Time operates in cycles, such that what has happened and worked in the past can and should inform our behavior in the present in order to maximize our chances at success or happiness in the future. Although Mikulan and Sinclair (2023) discuss temporalities of repetition anchored in chance, difference, and infinite possibilities, they also suggest that temporalities of repetition anchored in sameness have an enduring impact in education.

Reading the notions of time-as-arrow and time-as-repetition together raises a third provocation about time: risk and uncertainty. If time is one-way, singular, and cyclical, then it raises the question of whether the uncertainties and potential dangers of life can be managed and anticipated by a sufficient reading of the present. Many dominant approaches within statistics and data science seek to answer this question in the affirmative. In 2020, the American Statistical Association, with the endorsement of the National Council of Teachers of Mathematics, released the *PreK-12 Guidelines for Assessment and Instruction in Statistics Education II* (GAISE II, Bargagliotti et al., 2020). GAISE II has since become a prevailing standards document in statistics education, serving as a guide for reform within the United States and around the world. These reforms aim to direct curricular and instructional focus towards data literacies, data

science, and statistical reasoning and proficiencies within K-12 education. GAISE II begins with imagery of a "digital revolution" consisting of "readily accessible data" and "statistical methods and technological tools" that will enable students to "*make recommendations to manage pressing world issues*" (Bargagliotti et al., 2020, p. 1, emphasis added). These "pressing world issues" include "the COVID-19 global pandemic, a changing planet with extreme weather conditions, economic upturns and downturns, and important social issues such as the Black Lives Matter movement" (p. 1). Data literacies and data science are upheld as socially desired skills because by using data's predictive functions, people can understand and manage future risks of climate change and economic and social collapse.

The dominant language of statistics and data science is one of risk management and aversion. The idea of protecting ourselves from risk not only refers to risks generated by the outside world but also by risks associated with the practice of statistics and data science itself. The authors of GAISE II state that one of the crucial functions of data literacies education is to empower students to recognize "misleading graphical representations and limitations of data sets" (Bargagliotti et al., 2020, p. 7). In the absence of a rigorous statistical problem-solving process, "biases and misuses might emerge" (p. 12). Saba's statements about data as "efficient" and "understandable" produced and was productive of a desire to combat biases and misuses of knowledge. In a later interview about her data postcards, Bobble stated:

> If you use really clear graphs, some methods make understanding easier. Some make the understanding part of the story itself. I really think the approach or the medium is important because it depends on information foremost. Data storytelling depends on information foremost. So a mistake in the telling can lead to a mistake in the understanding.

Saba likewise talked about her concern around making a mistake during the data collection process of the data postcard activity. She stated, "I could have easily forgotten to record certain

pieces of the data, or I could have recorded more pieces of the data which would have skewed it." These words, alongside the quantized emotions of Bobble's Lego-built data physicalization and the grid-like structure of Saba's data postcard, enact the possibility for data to take on a risk mitigating role. The purpose of data is to guard against bias and misunderstanding and to protect ourselves from any disjuncture of the steady rhythm of time. Prediction and transparency of communication becomes values to which all data literate students should aspire.

The notion of risk emerged once again within Saba and Bobble's research project on COVID-19 masking. The global COVID-19 pandemic gave rise to a new awareness of risk and a global response to mitigate it. People were required to weigh the costs and benefits associated with activities such as masking, quarantining, sending children to school, riding a plane, or going to work. Risk highlighted differences of political opinion, differences in levels and forms of privilege, and differences around conditions of precarity and vulnerability. These energies were re-awakened in the questions that Saba and Bobble asked in their survey, where they sought to gauge students' COVID-19 masking preferences across different scenarios. They accompanied their questions with infographics that they had found on the Internet and modified for purposes of the survey. These infographics re-enact the decision-making process around COVID-19 risk mitigation. Students are asked to rehearse the choice that they had been making for the past several years—whether to mask or not. Masks evoke the question: will the COVID-19 pandemic ever truly go away, and what do I need to do to protect myself from this uncertain future?

Beck (1992) writes about the modern era being one of a "risk society." Risk in this case is the "systematic way of dealing with hazards and insecurities induced and introduced by modernization itself" (p. 21). Industrialization has created undesirable side effects such as pollution, and these side effects have manifested at a scale that has begun to undermine the

institutions of modern society (Beck, 1992; Van Loon, 2002). Guided by recent advancements in statistics, these side effects have become framed primarily in terms of probabilistically-defined risk (Beck, 1992). As these risks proliferate, they paradoxically become generated by their articulation. Risk does not exist on its own but instead becomes a construct of statistically-based scientific analysis and communication. As Van Loon (2002) states, "Without symbolic forms, risks are nothing" (p. 29, citing Beck, 2000; Van Loon, 2000).

Saba and Bobble's COVID-19 masking infographics were not merely representations of risk. They were the symbolic (and material) forms of risk itself. This is not to say if we as a society closed our eyes to COVID-19, then the problem would go away. Rather, the notion of COVID-19 *risk* is generated by its articulation in data-driven terms. When risks are articulated in this way, it changes the conditions of our own articulation, creating an environment that distorts our relationships to uncertainty and ourselves (Beck, 1992). Beck describes this as *individualization*, whereby the individual is removed from being-in-the-world and re-embedded into a social order that expects them to manage personal and social risks through the use of mediational technologies (Beck 1992; Van Loon, 2002). Risk intensifies the "shifting of the burden of responsibility for decision-making to the level of the individual. Do we eat beef or not? What are beef products anyway?" (Van Loon, 2002, p. 30). In a risk society, people need to develop the skills necessary to evaluate and guard against the hazards of an uncertain world. If a negative consequence befalls them, it is their responsibility to shoulder these consequences because they have failed to sufficiently protect themselves against risk.

The question of whether to wear a mask or not emerges alongside this notion of individualization. Protecting oneself against bodily harm and preventing the transmission of a potentially deadly virus becomes a matter of *choice*. The knots in Saba and Bobble's yarn-based

data physicalization are emblematic of these choices. They tell a story of individuals weighing the costs of present action against the potential costs of an uncertain future. These knots of yarn are entangled with discourses in statistics and data science education policy documents. The authors of GAISE II state, “Good data sense is needed to easily read the news and to participate in society as a well-informed member” (Bargagliotti et al., 2020, p. 1). The authors further discuss the value of statistics collected from personal fitness devices as “a way that motivates healthier lifestyles” as well as the value of data allowing people to “make decisions based on statistics about the cost of living or local climate” when they relocate to new communities (p. 5). Rather than asking students to question governmental and institutional failures to manage risks at a collective level through healthcare, economic, or climate regulation, GAISE II and other related texts assume that the solution to society’s future problems is to equip individuals with the information and skills required to participate within a datafied society. In this imagined future, individuals use data-related skills to make personal financial and health decisions, guard themselves against biased news sources, and use data-driven decisions to improve their quality of life.

These discourses promote the idea that there are high stakes associated with data that may not be present in other forms of knowledge production. When explaining how storytelling with data differed from other forms of storytelling, Bobble stated that making mistakes “is less important in other storytelling, like fiction. Like if you misunderstand something about a character, it doesn't matter, the person's not real, but if somebody misunderstands something or insulin that can have real actual issues in their life.” By evoking the potentially costly consequences of misunderstandings with data, Bobble enacted a boundary separation between what is “real” and “not real” and what is “reality” and what is “merely fiction.” Thus,

probabilistically-defined risk became not only a matter of future temporality but also that of the present insofar as it partitioned what should be regarded as "real" versus "not real" today, or what is important versus what is merely fiction.

Discussion and Implications

Despite my experience with and enthusiasm for socio-material theories—many of which eschew the post-positivism often promoted within data literacies, data science, and statistics—the course that I helped develop and teach was a space where many post-positivist and other traditional ideas about data emerged. The hands, gestures, materials, and discourses of the students, along with my own assignments and involvement in the course, made possible the idea, for instance, that data could be used to understand the world from a distant, disembodied, and detached perspective. Ironically, this "god trick" of data (D'Ignazio & Klein, 2020) emerged from data's own materiality, whether it was Saba pointing to separate elements of her Lego sub-structures to explain what they "represented" or whether it was Bobble quantizing her own emotional states and mapping them to colors in her Lego-built data physicalization. Instructional design and pedagogy, therefore, is not a matter of achieving a predetermined end goal that is planned in advance. Rather, educators have no choice but to move *through* the instructional designs they purport to construct from the outside. Instead of the designer being at the forefront of the design process, they are one of an array of human and more-than-human agencies who "come into being and are rendered capable through multidirectional relationships" (Bozalek & Zembylas, 2017, p. 64). The idea that "we are all part of the world, and that we cannot distance ourselves from it or assume a stance of innocence in our relationships with others" (p. 68) grounds what Bozalekak and Zembylas (2017) refer to as response-able pedagogies. Instructional design and pedagogy is not about designing interventions based on evidence-based results. It is

about taking up one's response-ability as a participant in a design process where all agencies are enacting boundaries and exclusions of some kind. It is through these boundaries and exclusions that possibilities for learning throughout the course emerged or became foreclosed. Put another way, one significant implication of this paper is that one cannot force one's vision of reform-minded education onto students. Although I had particular ideas about data in mind that were informed by critical and new materialist perspectives, traditional epistemological and ontological orientations to data continued to surface within my students' words and gestures and even within my own instructional practices. Moreover, even to the extent that students were willing or eager to take up ideas that draw on a relational ontology of data, these ideas may be so unfamiliar that students interpret them in ways that do not align with my own emerging understanding of these ideas.

Instead of seeking to impose knowledge onto students, part of a response-able orientation to pedagogy and instructional design might include closely attending to and having explicit conversations with students about the normativity of disciplinary knowledge and the reasons why one would want to think carefully about the nature of knowledge in the first place. Throughout my analysis, I discussed how students, teachers, and materials alike produced both traditional understandings of data and new onto-epistemic possibilities for data. These understandings of data, however, are not dichotomous and do not form a binary between what may be considered desirable and what may be considered undesirable for students to learn about data. The idea of data as providing a distant and detached perspective on the world, for instance, does not in itself suggest something inherently "good" or "bad." Once it is observed that agentic participants in a learning environment are creating possibilities for the "god trick" of data to emerge, it is not the case that such possibility be immediately resisted or promoted. D'Ignazio and Klein (2020)

approach their criticism of the "god trick" of data from the perspective of intersectional feminism and the interests and values that it advances. At the same time, scholars such as Wise (2020) and Konold et al. (2015) might claim that such distance is necessary in order for students to use data to understand phenomena from an aggregate perspective. Both understandings of data may be appropriate depending on one's particular context, including the backgrounds, needs, and interests of one's students. Thus, pedagogical and instructional design practices within data literacies, data science, and statistics education are a matter of attending to the various onto-epistemic assumptions about data that arise on account of the relational practices, affective intensities, and flows of discourse-materials that occur within the classroom. These practices, intensities, and flows give rise to an obligation of response, but these responsibilities emerge in the moment and through consideration of the entire learning context.

In practice, this suggests that one overlooked role of a teacher within a data-intensive learning environment is to attend to the normativity of disciplinary approaches to data that emerge within the space and to engage their students in explicit discussions about one's epistemological and ontological orientations to data. For instance, as I was working with Saba as she incorporated the cookie into her data physicalization, or as I was working with Bobble as she began to narrate a new data story while disassembling the Legos in her hand, I could have used these moments as an opportunity to begin a conversation with them and the entire class about what is regarded within various communities as a "legitimate" data representation. Indeed, such a discussion would be part of a broader ongoing debate about best practices with respect to data representation, with some scholars believing that data visualization should be minimalistic and devoid of superfluous aesthetic choices while other scholars promote embodiment and affect within such practices (D'Ignazio & Klein, 2020). Paying attention to the epistemological and

ontological orientations to data that emerge from the discursive-material intra-actions of the classroom, therefore, offers an opportunity to enter into existing conversations or spark new ones about the nature of knowledge itself. This is akin to the notion of "meta-commenting" discussed in Herbel-Eisenmann et al. (2009) and Chazan & Pimm (2016), but with emphasis placed not only on the impacts of student and teacher discourse on student learning and identity development but also on the agentic capacities of the more-than-human world and the entanglements between the material-discursive environment of the classroom and broader politics of disciplinary knowledge. Questions that teachers and students might consider include: What effect are particular onto-epistemic assumptions of data having on the students and on the instructor? What possibilities for learning are being promoted or foreclosed? How do these assumptions connect to one's larger goals for education? What if different onto-epistemic possibilities for data emerged, as in the case of possibilities that elevate the materiality of data or queer our expectations about clock time? In that case, whose bodies and modes of being become privileged and whose become marginalized?

Questions such as these suggest a potential shift away from teaching and instructional design as a scientific endeavor of formulating universal principles based on the empirical results of experiments. Instead, teaching is a matter of co-constructing knowledge with all participants in a learning space, where participants include not only the teacher and students but also the more-than-human world of desks, chairs, walls, pens, paper, cookies, and Legos that also shape the possibilities for learning, growth, change, and response that are the outcomes of any pedagogical encounter. Following Bozalek and Zembylas (2017), pedagogy and instructional design are a call to "pay due attention, to read, or listen with discernment and care" (p. 67), make way "for the possibility to be surprised and intrigued through unanticipated encounters" (p. 68),

recognize that "difference is embodied and entangled" (p. 69), and attend to the idea of "being rendered or rendering each other capable" (p. 69). These elements of Bozalek and Zembylas's framework of response-able pedagogy refer not only to how teachers and students should treat one another but how they can collectively approach learning about data together. Paying "due attention, to read, or listen with discernment and care" can refer not only to listening to what students are saying and doing but also to attuning to the intra-active entanglements of teachers, students, and materials and how such entanglements give rise to both traditional and novel ways of thinking about and using data. Data-intensive learning environments, moreover, should not be a space where settled "best" practices around data are passed down from instructor to students. Instead, data-intensive learning environments are spaces where novel ideas about data can emerge when teachers and students are willing to be "surprised and intrigued through unanticipated encounters" and when difference is recognized as "embodied and entangled." In the case of my conversations with Bobble and Saba, for instance, a novel question about data arose that I had not anticipated when I began this study: Is a more-than-representational data science possible? What might that look like? And although an answer to this question was only suggested by Bobble's act of disassembling the Legos or Saba's act of pulling in a cookie into her data physicalization, the mere raising of the question was itself an important moment that has powerful implications for how I might approach instructional design within data-intensive learning environments in the future. In short, data is not merely the passive "context" of a data-intensive learning. It is an active participant (see Chen, 2023) that plays a role in "being rendered or rendering each other capable" alongside teachers and students. This paper's analysis is an attempt to highlight the possibilities for teaching and learning that occur when one adopts such a stance about data and about knowledge broadly.

Conclusion

This paper elevates the importance of attending to relational practices and affective intensities within instructional designs for data literacies, data science, and statistics education. These practices and intensities not only included the words and actions of Saba and Bobble but also data assemblages made up of Legos, cookies, data postcards, book cycles, yarn, and discourses about materiality, time, and risk. I briefly return to Constance, whose words constituted the epigraph of this paper:

> I remember talking about how there's such a bigger picture with data and you read, oh, this percent of something. And it's like, okay, well that's like a cold, hard fact, but really no. It's like, who surveyed those people? Who were the people surveyed? What… why was… who were the people surveyed? There's so much story behind every single piece of data and you can't even fully take the data to be, oh, this is a fact, without really understanding where it comes from and who it comes from and who the data's about.

In this paper, I sought to theorize—and in doing so, specify—the "bigger picture" and the "story behind every single piece of data" that Constant refers to. I attuned to the agentic role that data assemblages play within the data-intensive learning environment of the data storytelling course. In doing so, I sought to more deeply understand how people *and* materials enacted boundaries and made certain experiences possible while excluding other potential experiences.

Diffraction was an important methodological approach toward my goal of analyzing the results of what I characterize as a teaching and learning experiment. My co-instructor and I developed the course from the ground up with no robust curriculum to guide us. Before the course began, my co-instructor raised the question of how we would know what the students learned and how well they learned what they learned. One possibility that was raised and subsequently developed was the use of individual and group interviews, along with the outcomes of student work from assignments and longer term projects. These methods of receiving

feedback within the instructional design process, however, are limited by their preoccupation with language. Many socio-material methodologies begin with the premise that "[l]anguage has been granted too much power" (Barad, 2003, p. 801). In order to de-privilege the status of language and gestures, I sought to consider how the Legos, postcards, and yarn co-constituted and had a performative effect on students' orientations toward the nature of data. Materials cannot be inert participants within classrooms because "all matter works together to co-produce meaning in the ongoing flow of experience and the entanglement of agencies" (Ehret et al., 2016, p. 348). Paying attention to the agentic capacities of materials tells a different and potentially richer story about teaching and learning, not only because it acknowledges previously unacknowledged agentic participants within the classroom but also because it acknowledges that classrooms are themselves materially rich, dynamic, and ever-shifting spaces. As Ehret and colleagues (2016) state, "[C]apacities to act…are enmeshed in ongoing, moving intra-activity, wherein bodies-materials' capacities to act shift across settings" (p. 352). Pedagogy therefore is a matter of attuning to what capacities are made possible and what capacities are foreclosed.

Course participants reproduced and disrupted ideas about what data are, their materiality, their relation to the world, their temporality, and their role in ideas about certainty and risk. These participants included not only Saba and Bobble but also the ocean of materials within which they swam, the storylines that emerged from the discourse-materials within the classroom, the data assemblages that co-constituted their learning activities in the data-rich environment, and broader assemblages of power. These participants also included me and my co-instructor. One implication of socio-material methodologies is that it is impossible to separate any observation from its corresponding act of observation; and likewise, it is impossible to separate any act of observation from its observer (e.g., Barad, 2007). While the course was replete with

moments where both traditional ideas about and new possibilities for data emerged, I came to know these moments through my own affective involvement in the process. This involvement included interviews that I conducted with a camcorder or phone. It also included my role as an instructor carrying the weight, influence, and apparitions of past teaching experiences along with my future aspirations as an emerging scholar.

Breaking free from the idea of learning design or pedagogy as the product of deliberate action grounded in static principles, this analysis suggests that teaching, learning, and design always involves a degree of *intra-ruption*. Throughout this paper, I attuned to how intra-ruption occurs across different aspects of the data assemblages that emerged from the myriad intra-actions within the course. First, intra-ruption came from data themselves. This occurred in the unanticipated responses that Saba and Bobble received in their survey responses, which this analysis has characterized as apparitions or "ghostly matters" (Dixon-Román, 2017, p. 47) that "contaminate[] the haunted products of social inquiry" (p. 46). Data literacies, data science, and statistics are often concerned with techniques to erase these ghostly matters or contain them in the form of confidence intervals and error bars. This socio-material analysis, however, shows that these apparitions cannot be fully erased. They were not only prevalent in Saba and Bobble's process of data analysis but could also be read in the fraying strands and fragile entanglements of their yarn-based data physicalization. Data were therefore not the only agentic sources of intra-ruption in the course. Materials and bodies intra-rupt all stages of the data analytic process. The cookie became an unexpected participant in Saba's Lego-built data physicalization, thereby contributing a sense of playfulness and celebration that added noise to her feelings of stress about her water consumption patterns. As this cookie came into play, it rendered more visible the entanglements among the data physicalization activity, the broader classroom, and the generosity

and joy present on that last day of class in the Fall semester. No instructional design principle could anticipate how Bobble's family's culinary hobby could find its way into Saba's data physicalization and eventually shape her narrative about her data postcard. Even the Legos themselves became a source of surprise, as they evoked in Saba memories of building Legos with her brother. Although the activity was focused on *building* Legos to create a data physicalization, Bobble unexpectedly used the *disassembly* of Legos to jumpstart a new idea about her data. Data is normally cleaned, curated, and built up in order to generate insight about a phenomena in the world. In contrast, this analysis shows that insight can emerge when we embrace data's production of excess and allow what was previously built up to crumble and fall apart. Dixon-Román (2017) states, "It is the excess, the beyond, the seeming erasure, or the mutually constituted Other of the referent that remains present as ghostly matter" (p. 56). Attending to these connections and entanglements illuminates the relational practices and affective intensities that can give rise to new possibilities for engagement within the classroom. Intra-ruptions is a particularly apt concept for data-intensive learning environments because, as this analysis shows, they are an integral and ever-present force within data practices and yet they are often ignored, taken for granted, and relegated to the realm of error and noise. In what ways do teachers, students, and designers permit the possibility for intra-ruptions to occur? What is excluded when intra-ruptions are prevented or curtailed for the sake of clarity and efficiency?

It is also worth noting the unanticipated composition of students who formed the course. Upon receiving the student roster before the first day of class, I was immediately struck by the fact that the course consisted entirely of students who identified as women. My co-instructor hypothesized that this may have occurred because the course description included a reference to *Data Feminism* as one of the course texts. Regardless, although the course was designed with the

goal of reimagining possibilities for data literacies and data science education before any students were registered for the course, it was fortuitous that engaging with data with an all-female class in a novel way through data physicalizations and storytelling practices afforded an opportunity to speak back against the patriarchal ways of knowing, being, and doing that typically characterize traditional approaches to data (D'Ignazio & Klein, 2020). Bobble's use of yarn, for instance, was striking because of the "long-standing association of craft with woman's work" (Myzelev, 2015, p. 62), and the "denigration of knitting correlates directly with the denigration of a traditionally women-centered activity" (Pentney, 2008, p. 1). Therefore, the use of yarn as part of a data physicalization could be interpreted as a feminist practice (Pentney, 2008). However, when my co-instructor and I asked Bobble in a follow-up interview about whether there was any gendered performance involved in the use of yarn in her and her group's data physicalization, Bobble indicated that she was not aware of this connection. Nonetheless, although there may not have been intentionality in her group's decision to use yarn as a feminist practice, I argue that the intra-ruptive force of the yarn—which included the possibility for telling a data-driven story in a humanizing way alongside the knots, tangles, and spills that constituted the yarn's agentic capacity to create an uncapturable "excess"—continued to carry the weight of a feminist practice in form of a "ghostly matter."

Lastly, attending to the role of intra-ruption within data-intensive learning environments points to two additional forces that contributed to the intra-active becomings within the course: improvisation and storytelling. In the case of Saba, her action of incorporating the cookie into her Lego-built data physicalization was a moment of material improvisation and storytelling. The cookie's presence was an important agentic force within the classroom, and so too was Saba's impulse to draw it closer to her Legos. For Bobble, the story of her book cycle emerged

spontaneously as she was disassembling her Lego-built data physicalization. In both instances, improvisation and storytelling were not the product of random impulse. They are practices that emerge from one's ongoing enmeshment and becoming within an ocean of materials. Fois (2022), in writing about musical and movement-based jamming through a new materialist lens, considers the possibility of jamming as "a material-discursive practice in which traditional notions of subject and object are questioned" (p. 240). Jamming is an open-ended improvisational play session involving human and more-than-human agencies that call on participants to listen attentively and resist the urge of seeing oneself as a fully formed identity. Through jamming and improvisation and "through the entanglement of our creativity, environment, and histories, we experience a change of self. The process of jamming is the process of being 're-formed' and always 're-forming'" (p. 248). The data postcard activity, including the follow up activities involving Legos, created the possibility for jamming and play with data, where the students were not masters over their materials but instead co-participants in the formation of data-driven ideas. These ideas were not reflections of an independently existing external reality but rather part of the process by which Saba and Bobble were continuously being formed and re-formed.

Thinking with notions of improvisation and storytelling, this analysis suggests that teaching, learning, and design for data-intensive learning environments cannot be reduced to a matter of pre-planning deliberate action, observing results, and forming evidence-based teaching and design principles. I have suggested a move toward a response-able pedagogy (Bozalek & Zembylas, 2017) that centers attentiveness, curiosity, and response. Data is not merely a tool for telling stories about a world that is outside of ourselves. It is a dynamic and open-ended assemblage of discourse-materials that interrupt and become interrupted by our engagements in

the world. It is a means of telling stories about a world of which we are an integral part. Data are moments of "self-touching [that] is an encounter with the infinite alterity of the self" (Barad, 2012, p. 213). Teaching, learning, and designing within data-intensive environments, then, are acts of storytelling not from the outside but from within stories themselves.

REFERENCES

Andres, L. (2012). *Designing and doing survey research*. SAGE Publishing.

Augustine, S. M. (2014). Living in a post-coding world: Analysis as assemblage. *Qualitative Inquiry, 20*(6), 747-753.

Bae, S. S., Szafir, D. A., & Do, E. Y. L. (2022). Exploring the benefits and challenges of data physicalization. *Proceedings of ETIS*.

Bahng, A. (2018). *Migrant futures: Decolonizing speculation in financial times*. Duke University Press.

Bang, M. (2017). Towards an ethic of decolonial trans-ontologies in sociocultural theories of learning and development. In I. Esmonde, & A. N. Booker (Eds.), *Power and privilege in the learning sciences: Critical and sociocultural theories of learning* (pp. 115-138). Routledge.

Bang, M., & Vossoughi, S. (2016). Participatory design research and educational justice: Studying learning and relations within social change making. *Cognition and instruction*, *34*(3), 173-193.

Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society, 28*(3), 801-831.

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.

Barad, K. (2010). Quantum entanglements and hauntological relations of inheritance: Dis/continuities, spacetime enfoldings, and justice-to-come. *Derrida Today, 3*(2), 240-268.

Barad, K. (2012). On touching—The inhuman that there I am. *differences: A Journal of Feminist Cultural Studies, 25*(3), 206-223.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D.A. (2020). *PreK-12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education*. American Statistical Association and National Council of Teachers of Mathematics.

Beck, U. (1992). *Risk society: Toward a new modernity*. Sage Publications Inc.

Beck, U. (2000). Risk society revisited: Theory, politics and research programmes. In B. Adam, U. Beck, & J. Van Loon (Eds.), *The risk society and beyond: Critical issues for social theory* (pp. 211-229). Sage.

Bhargava, R., Deahl, E., Letouzé, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). Beyond data literacy: reinventing community engagement and empowerment in the age of data.

Data-pop alliance white paper series. *Data-Pop Alliance White Paper Series*. Accessed at: https://datapopalliance.org/wp-content/uploads/2015/11/Beyond-Data-Literacy-2015.pdf

Bingham, C. W., & Sidorkin, A. M. (Eds.). (2004). *No education without relation*. Peter Lang.

Bogost, I. (2012). *Alien phenomenology, or what it's like to be a thing*. University of Minnesota Press.

Bogue, R. (2013). *Deleuze on literature*. Routledge.

Bozalek, V., & Zembylas, M. (2017). Diffraction or reflection? Sketching the contours of two methodologies in educational research. *International Journal of Qualitative Studies in Education, 30*(2), 111-127.

Branch, R. M., & Kopcha, T. J. (2014). Instructional design models. In J. Spector, M. Merrill, J. Elen, & M. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 77-87). Springer.

Brinkmann, S. (2014). Doing without data. *Qualitative Inquiry, 20*(6), 720-725.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry, 18*(1), 1-21.

Cairo, A. (2016). *The truthful art: Data, charts, and maps for communication*. New Riders.

Cajete, G. (2000). *Native science: Natural laws of interdependence.* Clear Light Publishers.

Chazan, D., & Pimm, D. (2016). Dilemmas and the teaching of mathematics: A conversation of commitments, obligations, and ambivalence. In M. Phakeng, & S. Lerman (Eds.), *Mathematics education in a context of inequity, poverty and language diversity* (pp. 19-31). Springer.

Chen, G. A. (2023). Teacherly response-ability: ethical relationality as protest against mathematical violence. *Educational Studies in Mathematics*, 1-20.

Cohen, P. C. (1982). *Calculating people: The spread of numeracy in early America*. University of Chicago Press.

Coole, D., & Frost, S. (2010). *New materialisms: Ontology, agency, and politics*. Duke University Press.

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking Big Data's relation to the contemporary subject. *Television & New Media, 20*(4), 336-349.

Crawley, A. T. (2016). *Blackpentecostal breath: The aesthetics of possibility*. Fordham University Press.

Data Science for Everyone (n. d.). About the coalition. Accessed at: https://www.datascience4everyone.org/about

Decuypere, M. (2019). STS in/as education: where do we stand and what is there (still) to gain? Some outlines for a future research agenda. *Discourse: Studies in the Cultural Politics of Education, 40*(1), 136-145.

de Freitas, E., & Sinclair, N. (2013). New materialist ontologies in mathematics education: The body in/of mathematics. *Education Studies in Mathematics, 83*, 453-470.

de Freitas, E., & Sinclair, N. (2014). *Mathematics and the body: Material entanglements in the classroom*. Cambridge University Press.

Deleuze, G., & Guattari, F. (1987). *A thousand plateaus: Capitalism and schizophrenia*. University of Minnesota Press.

D'Ignazio, C., & Klein, L. (2020). *Data feminism*. The MIT Press.

Dixon-Román, E. (2017). Toward a hauntology on data: On the sociopolitical forces of data assemblages. *Research in Education, 98*(1), 44-58.

Draper, E. B. (2020, December 16). Data science job growth in 2021 and beyond. *The Commons*. Accessed at: https://lambdaschool.com/the-commons/data-science-job-growth-in-2021-and-beyond

Duncheon, J. C., & Tierney, W. G. (2013). Changing conceptions of time: Implications for educational research and practice. *Review of Educational Research, 83*(2), 236-272.

Dunne, A., & Raby, F. (2013). Speculative everything: Design, fiction, and social dreaming. The MIT Press.

Ehret, C., Hollett, T., & Jocius, R. (2016). The matter of new media making: An intra-action analysis of adolescents making a digital book trailer. *Journal of Literacy Research, 48*(3), 346-377.

Ernest, P. (1991). *The philosophy of mathematics education*. Falmer.

Feigenbaum, A., & Alamalhodaei, A. (2020). *The data storytelling workbook*. Routledge.

Fois, L. (2022). A new materialist twist on the ontology of play. *American Journal of Play, 14*(3), 233-253.

Fotopoulou, A. (2021). Conceptualising critical data literacies for civil society organisations: Agency, care, and social responsibility. *Information, Communication & Society, 24*(11), 1640-1657.

Fox, N. J., & Alldred, P. (2014). New materialist social inquiry: Designs, methods and the research-assemblage. *International Journal of Social Research Methodology, 18*(4), 399-414.

Gill, M. G., Ashton, P. T., & Algina, J. (2004). Changing preservice teachers' epistemological beliefs about teaching and learning in mathematics: An intervention study. *Contemporary educational psychology, 29*(2), 164-185.

Gorur, R., Hamilton, M., Lundahl, C., & Sjödin, E. S. (2019). Politics by other means? STS and research in education. *Discourse: Studies in the Cultural Politics of Education, 40*(1), 1-15.

Green, K. L. (2020). Radical imagination and "otherwise possibilities" in qualitative research. *International Journal of Qualitative Studies in Education, 33*(1), 115-127.

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, *14*(3), 575–599.

Hashweh, M. Z. (1996). Effects of science teachers' epistemological beliefs in teaching. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 33*(1), 47-63.

Herbel-Eisenmann, B., Drake, C., & Cirillo, M. (2009). "Muddying the clear waters": Teachers' take-up of the linguistic idea of revoicing. *Teaching and Teacher Education*, *25*(2), 268-277.

Ingold, T. (2011). *Being alive: Essays on movement, knowledge and description.* Routledge.

Jackson, A. Y. (2013). Posthumanist data analysis of mangling practices. *International Journal of Qualitative Studies in Education, 26*(6), 741-748.

Jackson, A. Y., & Mazzei, L. (2011). *Thinking with theory in qualitative research: Viewing data across multiple perspectives*. Routledge.

Jefferson, T. (1954). *Notes on the state of Virginia*. Norton. (Original work published 1784)

Kahn, J. B., Peralta, L. M., Rubel, L H., Lim, V. Y., Jiang, S., & Herbel-Eisenmann, B. (2022). Notice, wonder, feel, act, and reimagine as a path toward social justice in data science education. *Educational Technology & Society, 4*(25), 80-92.

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures & their consequences*. Sage.

Klykken, F. H. (2021). 'Are *we* going to do *that now*?' Orientations and response-abilities in the embodied classroom. *Pedagogy, Culture & Society, 1-17*.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics, 88*(3), 305-325.

Le Blanc, A. M. (2017). Disruptive meaning-making: Qualitative data analysis software and postmodern pastiche. *Qualitative Inquiry, 23*(10), pp. 789-798.

Lee, V. R., Wilkerson, M. H., & Lanouette, K. (2021). A call for a humanistic stance toward K–12 data science education. *Educational Researcher, 50*(9), 664-672.

Lim, V., Peralta, L. M., Rubel, H., Jiang, S., Kahn, J. B., & Herbel-Eisenmann, B. (2022). Keeping pace with innovations in data visualizations: A commentary for mathematics education in times of crisis. *ZDM, 55*, 109-118.

Loukissas, Y. A. (2019). *All data are local: Thinking critically in a data-driven society*. MIT Press.

Lupi, G., & Posavec, S. (2016). *Dear data*. Princeton Architectural Press.

Massey, D. (2005). *For space*. Sage.

McQuillan, D. (2018). Data science as machinic neoplatonism. *Philosophy & Technology, 31*(2), 253-272.

Mikulan, P., & Sinclair, N. (2023). *Time and education*. Bloomsbury.

Molenda, M., Reigeluth, C.M. and Nelson, L.M. (2006). Instructional design. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (pp. 574-578). Nature Publishing Group.

Myzelev, A. (2015). Creating digital materiality: Third-wave feminism, public art, and yarn bombing. *Material Culture, 47*(1), 58-78.

Oppermann, S. (2021). New materialism and the nonhuman story. In J. Cohen & S. Foote (Eds.), *The Cambridge companion to environmental humanities* (pp. 258-272). Cambridge University Press.

Pentney, B. A. (2008). Feminism, activism, and knitting: Are the fibre arts a viable mode for feminist political action?. *Thirdspace: a journal of feminist theory & culture*, *8*(1), 1-17.

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction, 34*(4), 361-388.

Pickering, A. (1995). *The mangle of practice: Time, agency and science*. University of Chicago Press.

Richardson, L., & St. Pierre, E. A. (2005). Writing: A method of inquiry. In N. K. Denziin, & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 959-978). Sage Publications Ltd.

Rifkin, M. (2017). *Beyond settler time: Temporal sovereignty and Indigenous self-determination*. Duke University Press.

Rubel, L. H., Herbel-Eisenmann, B., Peralta, L. M., Lim, V., Jiang, S., & Kahn, J. B. (2022). Intersectional feminism to reenvision mathematical literacies & precarity. *Research in Mathematics Education, 24*(2), 224-248.

Sabin Center for Climate Change Law (n. d.). COVID-19 data misrepresented by Georgia health department. Accessed at: https://climate.law.columbia.edu/content/covid-19-data-misrepresented-georgia-health-department

Serres, M. (1982). *The parasite*. The Johns Hopkins University Press.

Smith, A. (1937). *An inquiry into the nature and causes of the wealth of nations* (5th ed.). Random House. (Original work published 1789)

Taguchi, H. L. (2012). A diffractive and Deleuzian approach to analysing interview data. *Feminist Theory, 13*(3), 265-281.

Tate, W. F. (1995). Returning to the root: A culturally relevant approach to mathematics pedagogy. *Theory Into Practice, 34*(3), 166-173.

Thompson, M. (2017). *Beyond unwanted sound: Noise, affect and aesthetic moralism*. Bloomsbury Publishing USA.

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry, 16*(10), 837-851.

Tsai, C. C. (2000). Relationships between student scientific epistemological beliefs and perceptions of constructivist learning environments. *Educational Research, 42*(2), 193-205.

Vakil, S., McKinney de Royston, M., Suad Nasir, N. I., & Kirshner, B. (2016). Rethinking race and power in design-based research: Reflections from the field. *Cognition and Instruction*, *34*(3), 194-209.

van der Tuin, I., & Dolphijn, R. (2012). *New materialism: Interviews & cartographies*. Open Humanities Press.

Van Loon, J. (2000). Virtual risks in an age of cybernetic reproduction. In B. Adam, U. Beck, & J. Van Loon (Eds.), *The risk society and beyond: Critical issues for social theory* (pp. 165-182). Sage.

Van Loon, J. (2002). Risk and technological culture: Towards a sociology of virulence. Routledge.

Vossoughi, S., & Gutiérrez, K. D. (2017). Critical pedagogy and sociocultural theory. In I. Esmonde, & A. N. Booker (Eds.), *Power and privilege in the learning sciences: Critical and sociocultural theories of learning* (pp. 139-161). Routledge.

Wilkerson, M. H., & Laina, V. (2018). Middle school students' reasoning about data and context through storytelling with repurposed local data. *ZDM*, 50, 1223-1235.

Wise, A. F. (2020). Educating data scientists and data literate citizens for a new generation of data. *Journal of the Learning Sciences, 29*(1), 1-17.

CHAPTER 4: AN INTERDISCIPLINARY PERSPECTIVE ON THE DATA AND STORYTELLING COURSE

Introduction

There is no shortage of contexts within which data and data practices have shaped the way the world works. It has been estimated that by the end of 2020, there would have been as many bits of data in the digital universe as there are stars in the physical universe (Messy Data Coalition, 2020). The rapid growth in available data and the applications toward which data are used has given rise to metaphors of "data as the new oil," a "data as a revolution," or "data as a deluge" (Watson, 2015). These metaphors suggest a range of values and goals that social actors have attributed to data, including promises of profit from data analytics and big data, the use of data to solve pressing social problems, or the problems that arise from an oversaturation of data in today's society.

Regardless of the metaphor that one uses to describe data, across personal, social, and professional contexts students will need to produce, make sense of, critique, and communicate ideas based on data (e.g., Draper, 2020; NASM, 2018). This recognition has led to calls to expand opportunities for students to learn about data. These opportunities often manifest under the labels of data literacies and data science (hereafter, "data literacies/science"). Programs in data literacies/science have been most evident at the post-secondary level, where college course offerings, bootcamps, and online courses have proliferated to meet the surge in demand for data-related skills (Bhargava et al., 2015). The terms data literacies and data science do not have widely agreed upon definitions, and how these terms are operationalized tends to reflect the contexts in which they are used (Wolff et al., 2016). Data literacies/science skills are commonly associated with know-how around the collection, analysis, and representation of data (e.g.,

Bargagliotti et al., 2020). For a rapidly growing group of scholars, however, data literacies/science education should include more than the technical dimensions of these skills. These scholars call for supplementing technical proficiency with a critical and creative orientation toward data (Bhargava et al., 2015; D'Ignazio & Klein, 2020; Lee et al., 2021). This includes raising epistemological and ontological questions about the nature of data (Lee et al., 2021; Loukissas, 2019), attending to issues of power and ethics with respect to data (D'Ignazio & Klein, 2020), and thinking carefully about data's rhetorical and performative functions (Owens, 2011).

In this paper, we share results from a study on undergraduate students' understandings of data and perceptions of social issues through the lens of data. The study was a collaboration among nine undergraduate students, one graduate student, and one university professor in an interdisciplinary course on data and data storytelling at a large public university in the U.S. Midwest. The study combines a digital survey—distributed to a select group of undergraduate students—and corresponding data stories that the students developed to make sense of the survey results and learn about the process of data-driven research as a whole. The students chose the survey topics and questions, which deal with three issues of social concern: food insecurity, healthcare, and economic mobility. The study provides insights into how undergraduate students make sense of data in personally meaningful, real-world contexts.

The idea of data storytelling is of particular interest. We foreground data storytelling not only as a means to communicate data-driven findings in engaging and often novel ways, but also as a pedagogical tool for helping students develop critical and creative approaches toward understanding data's sociopolitical and material dimensions. As such, our study addresses the following research question: How did opportunities for critical, creative, and interdisciplinary

engagements with data throughout the data storytelling course shape how students made sense of data and processes of data generation, analysis, and communication? The paper proceeds in three parts. First, we discuss the relevant literature on data, data literacies/science education, and data storytelling. Then, we discuss the context of the study. Next, we discuss the methodologies of participatory research, abductive analysis, and voice-centered relational reading that form the basis of the study. Finally, we discuss our findings and the implications of our work for developing a critical and creative orientation to the teaching and learning of data among undergraduate students.

Literature Review

In a report on undergraduate data science education, the National Academies of Science and Medicine (NASM, 2018) argue that the "ability to measure, understand, and react to large quantities of complex data can shape scientific discovery, social interaction, political interactions and institutions, economic practice, public health, and many other areas" (p. 12). The perceived importance of data throughout all parts of society suggests the need to understand what exactly is meant by the concept of data in the first place. "Data" is broad and has no agreed upon meaning (Mertala, 2020), but some uses of the term closely associate it with measurements or sensor stimuli (Sanches et al., 2022) that are often, but not always, quantified and that are generated, stored, and translated across various media such as digital technologies (Loukissas, 2019). Others highlight data's close ties to statistics and characterize it as numbers in context (Cobb & Moore, 1997), as well as "dynamic, complex, highly structured (or unstructured) collections of pictures or sounds" (Bargagliotti et al., 2020, p. 6), which raises questions about who collected the data and when, where, why, and how such data were collected (Rubin, 2020).

The goal of data literacies/science can be said to create learning opportunities for students to develop "data acumen" (NASM, 2018, p. 22). The National Academies of Science and Medicine (NASM, 2018) organize the broad notion of data acumen into smaller parts that roughly correspond to disciplinary boundaries: mathematical foundations, computational foundations, data management and curation, data description and visualization, data modeling and assessment, workflow and reproducibility, communication and teamwork, domain-specific considerations, and ethical problem solving. These varied concepts reflect the interdisciplinary nature of data literacies/science education, raising both opportunities and challenges for the field. Because of the interdisciplinarity required to develop, understand, and communicate knowledge using data, the NASM (2018) report recommends a "unified approach" to data literacies/science that communicates "not simply a collection of varied tools (or methods), but rather a general approach to problem solving" (p. 20). This general approach to problem solving can include attending to context, variability, aggregate trends, visualization and other representational forms, and inferential reasoning with respect to a particular dataset (Rubin, 2020). While engaging in data-intensive activities, students might conceptualize data as a pointer to an event ("we said our favorite colors"), a case value that provides the value of an attribute of an individual case ("Juan likes red"), a classifier that provides frequency information ("three like red"), or an aggregate that provides information about an emergent property ("half like red") (Konold et al., 2015, pp. 308-309). The capacity to move between various perspectives on data is essential because "different contexts or questions may call for or cue different views of data even by the same student" (Konold et al., 2015, p. 309). The idea of a "unified approach" to data literacies/science education suggests the need and a starting point for a high-level roadmap for an interdisciplinary

data literacies/science curriculum that goes beyond teaching about data and data processes as a set of disjoint and discrete skills.

Even when one takes a high-level approach to data literacies/science curricular development, there is a risk that such an approach could promote an overly technical understanding of what engaging with data entails. The idea of data literacy as a purely technical pursuit is insufficient because doing so "reinforces existing inequities and should be replaced by the larger concept of inclusion" (Bhargava et al., 2015, p. 2), "disregards the need to address deeper structural issues of inequality" (Fotopoulou, 2021, p. 1641), and "help[s] students think critically about information and data, but not the structural and ideological contexts in which those information and data are embedded with meaning" (Philip et al., 2016, p. 365). Taking the view of data literacies/science as "not only a technical but also a sociocultural process" (p. 1641), Fotopoulou (2021) advances a five part notion of data literacies as agentic, contextual, critical, multiple, and social. These elements emphasize student agency in context-situated decision-making within data-driven processes, the ideological and power-laden aspects of working with and using data, and the connections that can and should be made between data usage and the collective goals of communities. Various studies in data literacies/science education reflect a commitment to promoting a view of data as emerging from and shaping sociocultural processes (e.g., Fotopoulou, 2021; Hardy et al., 2020; Kahn, 2020; Kahn et al., 2022; Lee et al., 2021; Lee & Dubovi, 2020; Philip et al., 2016; Rubel et al., 2021; Stornaiuolo, 2020; Van Wart et al., 2020; Wilkerson & Laina, 2018). In these studies, opportunities to engage with data in authentic, personally meaningful contexts enabled students to see themselves as agentic data producers and not merely collectors (Hardy et al., 2020) or consumers (Stornaiuolo, 2020) of data. Instead of imposing predetermined data analytic techniques on students based on

their own disciplinary backgrounds (Hardy et al., 2020), many of the educators in these studies allowed students make sense of the purposes and goals of their engagements with data (Fotopoulou, 2021; Van Wart et al., 2020). These purposes and goals included presenting data-driven findings to relevant authority figures (Van Wart et al., 2020), finding and accessing open data for use in advocacy and campaigning (Fotopoulou, 2021), using data to assist in family-based child care and management routines (Lee & Dubovi, 2020), and telling stories about oneself and one's family histories using georeferenced data (Kahn, 2020).

Technical data acumen, in other words, must be paired with a critical approach to data literacies/science that not only acknowledges but also deeply engages with its sociopolitical dimensions. Data are neither objective nor politically neutral despite impressions that data-savvy individuals are capable of understanding objects of study from a detached, distant, disinterested, and disembodied position (D'Ignazio & Klein, 2020). D'Ignazio and Klein (2020) call for approaches to data literacies/science that attend to issues of power, ethics, and justice throughout all aspects of the data-driven process. They argue that as part of resisting the tendency to see the data analyst as a detached, distant, and disinterested observer of phenomena, people should learn and embrace approaches to data literacies/science that acknowledge the relevance of emotions, the body, personal experience, and broader sociopolitical forces in shaping and being shaped by processes of data collection, analysis, and representation. One example of an approach to data literacies/science that connects personal experience with data is the Dear Data Project (Lupi & Posavec, 2016). The Dear Data Project was an arts-based project created by two briefly acquainted information designers who wanted to know how much they could learn about one another through sharing one another's data. The designers kept track of personally meaningful data about themselves for each week of one year, created a novel visualization based on their

data from that week, and sent each other the visualization via postcards. The activity was modified and used to productive effect with students in a study by Stornaiuolo (2020), which examined the data practices of young people in urban schools as they engaged in activities at the intersection of data and the arts. Pre-survey responses indicated that students saw data as something that was externally generated. As a result of the activities, the students began to report greater agency over data processes. This included identity development as producers and not mere consumers of data and understanding the close connection between data and the various purposes toward which data collection, analysis, and representation could be directed. Although the students used data to understand themselves in new ways and engage in the aesthetic dimensions of data representation, Stornaiuolo (2020) noted that these results tended to take precedence over interrogations over broader issues of the sociopolitical and cultural dimensions of data, which was another goal of the study design.

Other studies seeking to go beyond promoting an exclusive emphasis on technical data acumen examine the affordances of adopting a storytelling lens with respect to data literacies/science. Good data-driven scientific work involves good storytelling, which "recounts an expected cause that creates an effect for which we already have an explanation" (Matei & Hunter, 2021, p. 312). Good storytelling rests on the notion of surprise and requires a subversion of expectations and the provocation of new questions for further investigation (Matei & Hunter, 2021). This idea of data storytelling is important for data literacies/science education because "surprise begets questions, which begets interests, which begets personal engagement, which leads to learning" (Matei & Hunter, 2021, p. 314). Moreover, one study suggests that data can be understood *as* a form of storytelling. In particular, Wilkerson and Laina (2018), while writing about data-driven reasoning among middle school students, elaborate on the pedagogical benefits

of asking students to tell stories about data. Stories "can help make explicit the connections between the descriptive and inferential statistical methods usually kept separate in the curriculum," "help learners integrate new knowledge with what they already know," and "highlight the practice of data analysis as a messy, human endeavor" (p. 1224). Storytelling as an activity for data literacies/science education and as a frame for sense-making with data is an area ripe for further investigation.

The interdisciplinary course on data and storytelling about which this study is based was the result of attempts by Author 1 and Author 7 (the "instructional researchers") at considering, responding to, and taking up the ideas presented in the aforementioned literature. Regarding the day-to-day implementation of any course, however, there is rarely a seamless translation between the ideas that are presented within the literature and what occurs in the everyday reality of teaching and learning. The remainder of this paper discusses the interdisciplinary course in greater detail, as well as a discussion of the opportunities and challenges that emerged with respect to the goal of teaching and learning about data in deep and critical ways.

Context

This study is based on an interdisciplinary course on data and storytelling that the instructional researchers (Authors 1 and 7) proposed, designed, and taught through a university Honors College. Author 1 is a cisgender, straight Filipino-American male and doctoral candidate in curriculum and instruction with a former background in teaching secondary mathematics within public schools in New York City. Author 7 is a cisgender, white female and professor of Sociology within an interdisciplinary residential college housed within a major research university in the U.S. Midwest. The interdisciplinary course was a year-long research seminar that engaged students with various aspects of data storytelling. Nine first and second-year

students enrolled in the course, and all the students agreed to participate in the study. Their demographic information and pseudonyms are found in Table 2. Some of the students selected their own pseudonyms while others requested that Authors 1 and 7 choose their pseudonyms for them. The positionality of Authors 2, 3, 4, 5, and 6 are not shared in this paper in order to respect their use of pseudonyms below.

Table 2. Data Storytelling Course Participants

Name	**Gender Identification**	**Race/Ethnicity Identification**	**Year**	**Known Major/Area of Study**
Constance	Female	White	1st	Psychology
Dawn	Female	White	1st	Business
Adele	Female	White	1st	Computer Science
Eve	Female	White	1st	Communication Arts
Saba	Female	South Asian	1st	Pre-Medicine
Frodo	Female	White	1st	International Relations
Bobble	Female	White	2nd	Public Relations
Sasha	Female	Middle Eastern / White	1st	English/Psychology
Daisy	Female	Asian	2nd	Pre-Medicine

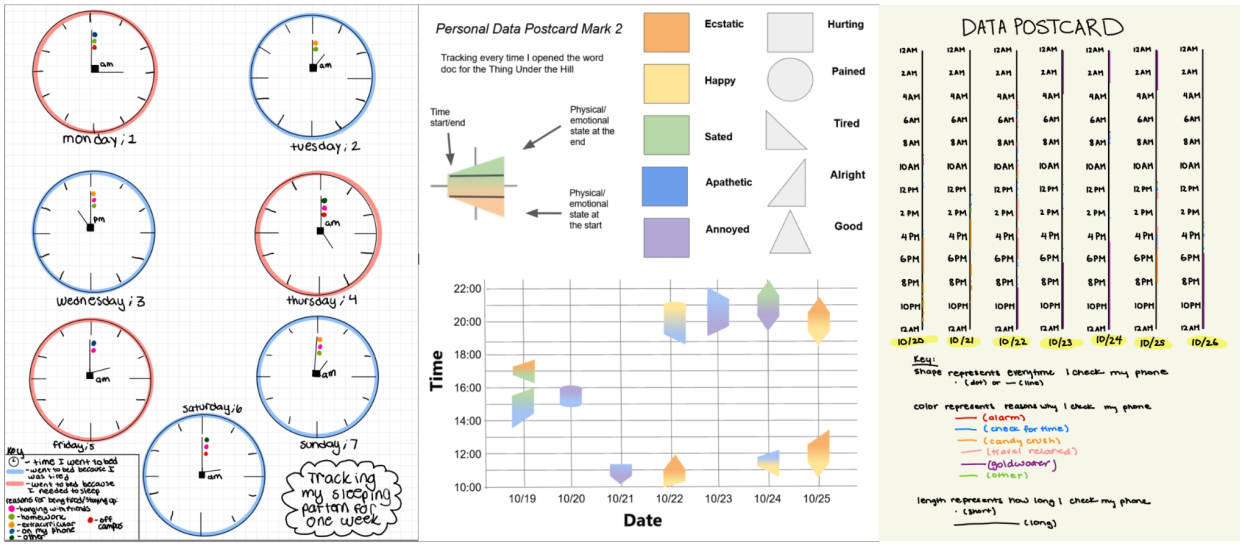
The year-long data storytelling course can be conceptualized in two parts: (1) a series of lessons about data collection, analysis, and representation, and (2) a significant student-led research project. In the first part of the course, which occurred in the first semester of the year-long course, students engaged with readings, discussions, and activities connected to various definitions, conceptualizations, and critiques of data and data literacies/science (D'Ignazio & Klein, 2020; Owens, 2011; Watson, 2015). The course design placed particular emphasis on identifying, analyzing, and crafting data-driven stories and understanding data from a storytelling

lens (Cairo, 2016; Feigenbaum & Alamalhodaei, 2020; Lupi & Posavec, 2018; Rakotondravony, 2019). One significant recurring assignment consisted of data story explorations, which asked students to search for, identify, and analyze news articles on the Internet that used a data representation as a storytelling device. In this assignment, as well as throughout the course, the instructional researchers (Authors 1 and 7) deliberately left the definition of what constitutes a "story" or what counts as "storytelling" open-ended in order to give students flexibility to interpret these terms in different ways. This created opportunities for class discussions about the boundaries of data storytelling and the elements or characteristics of effective and ineffective data-driven stories.

In addition, significant time was spent engaging students in iterative cycles of data generation and communication using the Dear Data Project (Lupi & Posavec, 2015), as discussed in this paper's literature review. The instructional researchers (Authors 1 and 7) modified the Dear Data Project and shaped it into a lesson that engaged students in creating analog representations of their own personally collected data. Such adaptation was inspired by and consistent with a number of freely available sample lessons on the Internet that make use of the Dear Data project in various instructional contexts (e.g., Quigley, 2020). In the version of the Dear Data project that the instructional researchers used for the data storytelling course, they asked students to collect data about one aspect of themselves over a seven day period and to create a hand-drawn representation of their data. They produced written reflections about the data collection process, what the data revealed about their lives, and how the process of producing their data postcard made them feel. They then re-created their data postcards using a digital technology and wrote reflections on whether changing the medium of representation changed their experience of designing their postcard, along with where variation appeared within

their data. Lastly, the students re-created their data postcards a third time in the form of a data physicalization using Legos. A selection of student work from the data postcard activity are included in Figure 10.

Figure 10. Select Data Postcards



The second part of the course, which began at the same time as the first part but extended into the second semester of the year-long course, involved the design of a survey disseminated to undergraduate students at the same university and the creation of data-driven stories based on the survey results. The students organized around three groups of three students each, with each group focusing on one of three social issues: food insecurity, economic mobility, and COVID-19 masking. The groups and topics were selected by the students, who were presented with a list of eight potential social topics (voting rights, climate justice, healthcare, refugee crisis, racial injustice, income gap, gun violence, and food insecurity) chosen by the instructional researchers (Authors 1 and 7). The instructional researchers developed these topics as a starting point and directed the students to modify or specify the topics as they chose. One group narrowed the broad topic of healthcare to the specific topic of COVID-19 masking. Another group changed the

topic of income gap to economic mobility. The last group kept the topic of food insecurity unchanged.

Due to time and institutional constraints, the work of the three groups had to be consolidated into a single survey that was then distributed to a sample of undergraduate students enrolled in interdisciplinary programs within the university. The survey-based project centered two research questions, which were co-developed by the instructional researchers and students: (1) How do undergraduate students respond to data artifacts and questions connected to contemporary social issues dealing with food insecurity, healthcare, and economic mobility? (2) How might the creation of critical and creative data-driven stories provide insights into how students make sense of data from an interdisciplinary perspective? Informed consent was obtained for both the research study within which the students were participants, as well as the survey-based study within which the students were researchers. Five hundred thirty two students responded to the survey using a Qualtrics survey disseminated via university emails. The course met for once a week in Fall 2022 with the expectation that research related work would continue in Spring 2023. The research experience culminated in three separate poster-based presentations that the students prepared and gave at an undergraduate research conference in April 2023.

Methodology

This study's methodological approach combines elements of participatory data analysis (Bourke, 2009; Bryne et al., 2009; Holland et al., 2008), abductive analysis (Brinkmann, 2014), and a voice-centered relational approach to reading student work and interview data (Paliadelis & Cruickshank, 2008; Letvak, 2003; Pinto, 2004). This section will begin by describing the process of data generation relevant to this paper. Then, we will discuss our process of data analysis in greater detail. In some qualitative studies (see Holland et al., 2008), as is the case

here, data generation and data analysis cannot be easily separated from one another, as both occur in mutually interdependent iterative cycles. Nonetheless, for purposes of this paper, we have separated data generation and data analysis into separate sections to facilitate readability.

Data Generation

It is important to clarify, at the outset, that this study consists of one study within another study. The innermost study consisted of the student research project described in the previous section. The outermost study, which is the subject of this paper, pertains to the research question of what undergraduate students learned about data through their experiences within the interdisciplinary course on data storytelling. Institutional review board approvals and participant consent were granted for both studies. The instructional researchers (Authors 1 and 7) informed the students about the research project on the first day of the data storytelling course. They solicited and received feedback from the students on the overall purpose and structure of the study. There was an acknowledged risk that the instructor-student relationship would cause students to feel compelled to participate or change their behavior throughout the course under the belief that non-participation or non-compliance would negatively impact their grades and the instructors' attitudes toward them. To minimize this risk, the instructional researchers asked a third party individual to obtain written consent from the students and withhold the details of the consent process from the instructional researchers until the course ended at the end of the year. Each of the nine students consented to participate in the study.

The instructional researchers held three video-recorded focal group interviews with the students throughout the Fall semester. These interviews took place at the end of select class sessions with most or all of the students present and were based on interview protocols developed by the instructional researchers using field notes and observations, which the

instructional researchers wrote after each class session. The first focal group interview asked students to describe their prior experiences with data and how the course had shaped their developing ideas about the nature of data. The second focal group interview asked the students about how the course had shaped their understanding of the relationship between data and power. The third focal group interview asked students to discuss how their thinking about the relationship between data and stories changed as a result of the course, along with any changes that they would recommend about the course structure. The instructional researchers chose these topics to highlight student learning throughout the course. These topics were also influenced by the instructional researchers' evolving teaching and research interests about data literacies/science education.

The instructional researchers also conducted individual interviews during select classroom activities, collected a wide range of student work, and recorded each class session when the course modality moved to Zoom during the second half of the course in the Spring semester as agreed upon by the instructional researchers and students in order to accommodate various scheduling conflicts that arose during the year. Lastly, the instructional researchers conducted one-hour follow up interviews with seven of the nine students during the summer after the course had ended. The individual interviews did not consist of an interview protocol that was strictly adhered to. Our aim was to conduct these interviews in the style of a conversation, which enabled participants to describe their experiences at length and to raise issues that the instructional researchers had not anticipated. Following each interview, the instructional researchers briefly discussed their thoughts and feelings about what the student and researchers discussed during the interview. Despite the relatively unstructured nature of the individual interviews, many of the interviews shared common themes raised by the instructional

researchers. In particular, the instructional researchers asked the students to discuss any information about themselves (e.g., race, gender, socioeconomic status, among other identity markers) that they believed to be relevant to their participation in the course. They also shared their thoughts, feelings, and recollections of three key activities in the course: the creation and analysis of their data postcards, the creation and analysis of their survey-based projects, and the presentation of their work at the undergraduate research conference.

Data Analysis

This study combines elements of participatory research (Bourke, 2009; Bryne et al., 2009; Holland et al., 2008), abductive analysis (Brinkmann, 2014), and a voice-centered relational approach to reading student work and interview data (Paliadelis & Cruickshank, 2008; Letvak, 2003; Pinto, 2004). Participatory research refers to processes that involve participants in decision-making around research design and implementation, including data generation, data analysis, and the dissemination of research results in public settings (Bourke, 2009). There are no strict rules about what constitutes participatory research (Bourke, 2009), and participatory research can take many forms across spectra of "informal and formal, unstructured and structured, trained and untrained, explicit and implicit" (Nind, 2011, p. 359) engagement. One's priorities in conducting participatory research may vary from instantiating a political commitment to inclusivity to striving for greater validity and insight in research results, among other non-mutually exclusive goals (Nind, 2011). Some scholars engage in participatory research in order to honor participants' authentic voices while others caution against interpretations of participant voices as inherently authentic, instead framing all voices as co-produced within dominant and non-dominant discourses (Nind, 2008). Participatory research also entails

analytical and ethical challenges raised by involving participants in decision-making around research processes and results (Bryne et al., 2009).

For purposes of this paper, we take the position that the voices of this study's student and instructional researchers are not inherently authentic because all voices and experiences are relational, in a constant process of becoming, and necessarily incomplete (Holland et al., 2008; Jackson & Mazzei, 2008). Our purpose for engaging participatory research, then, is to honor all of our incomplete voices—those of the instructors and students alike—as they express our developing understandings of teaching and learning within the data storytelling course. In the spectrum of different forms of participatory research, our approach sits between informal and formal and unstructured and structured. From the perspective of the instructional researchers, the form and purposes of participation were informed by several considerations and factors. First, while the student participants in the course were able to plan, develop, execute, and present their research projects at an undergraduate research conference, it was also discussed throughout the course that they would have the opportunity to co-author an academic paper if they desired. Therefore, we chose a participatory methodology in preparation of this paper in part to follow through on this plan. Second, much of the data analysis and co-writing occurred during a busy time of the academic year. Our methodological approach offered a wide range of participatory opportunities for student researchers while honoring students' limited time and energy and also showcasing alternative forms of research based on principles of relationality and self-awareness in research. The instructional researchers were also interested in participatory methods as a way to continue to mentor the students and carry out their evolving commitments towards combining research practices and pedagogy.

Shortly after the end of the course, the instructional researchers began the work of crafting this paper based on the aforementioned research questions on student learning about data. An initial meeting between the instructional researchers consisted of refining the research questions of the study and discussing a tentative approach to engaging the students in the preparation of the manuscript. The instructional researchers then reached out to the course participants to formally invite them to serve as co-authors to this paper. Five of the nine students agreed to participate (Authors 2, 3, 4, 5, and 6, hereafter "student researchers"). In addition, one of the instructional researchers and lead author of this article, who had prepared much of this paper's introductory remarks and literature review, received the relevant permissions to use the preparation of the manuscript as part of his dissertation study.

In their capacity as co-authors, the student researchers assisted in refining the research questions, engaging in data analysis, and writing and revising this paper's findings and implications. To facilitate this process, the instructional researchers (Authors 1 and 7) organized and held whole group and individual meetings where the instructors and students could discuss the paper and share general thoughts about its direction. Together, the student and instructional researchers identified potential sources of data to analyze. Based on this data, each student researcher prepared writing in their own time; the instructional researchers then provided comments and asked follow-up questions through emails with all of the student and instructional researchers copied on the messages, and the lead author helped synthesize the analysis and writing into this single manuscript. The instructional researchers shared drafts of this manuscript with the student researchers during iterative cycles of revision and feedback.

Participatory research, particularly when it involves participants in data analysis and manuscript preparation, suggests the opportunity and, at times, the need for methodological

flexibility and innovation (Byrne et al., 2007; Holland et al., 2008). Accordingly, the instructional researchers promoted a methodological approach to data analysis that included an orientation toward abductive inquiry (Brinkmann, 2014). Abduction is a form of inquiry focused on "situations of breakdown, surprise, bewilderment, or wonder" (Brinkmann, 2014, p. 722). Brinkmann contrasts abductive inquiry with induction, which seeks to generalize from particulars, and deduction, in which hypotheses are generated from theory. Central to abductive inquiry is the idea of "stumbling upon" (Brinkmann, 2014, p. 724) data by paying attention to moments that raise strong reactions on the part of the researchers. The researcher is an active part of the research process such that the goal of abductive inquiry is not to presume that a researcher can study an issue from a detached and disinterested vantage point (see Brinkmann, 2014). For this reason, abductive inquiry is a promising research orientation to pair with participatory methodologies because both enable participants to offer ideas informed by varied disciplinary backgrounds and positionalities. In other words, participatory methodologies offer a potentially powerful setting for facilitating the kind of breakdown and surprise necessary for successful abductive inquiry, and abductive inquiry offers methodological flexibility to elevate the voices of all researchers within a participatory setting.

We paired our process of abductive inquiry with elements of a voice-centered relational method (VCR) (Paliadelis & Cruickshank, 2008; Letvak, 2003; Pinto, 2004). VCR is an analytic approach to reading participant data grounded in a feminist orientation to knowledge production (Paliadelis et al., 2008). In VCR, the same instance of data is read multiple times based on a variety of perspectives. Brown and Gilligan (1992) systematized VCR into a four stage analysis, which we paraphrase based on our interpretation of the relevant literature: (1) What is the story and who is speaking? (2) In whose body and/or from what positionality? (3) What are the

relationships evoked by the story? and (4) What are the sociopolitical and cultural processes shaping or shaped by the story? VCR is particularly suitable for participatory research because it can "be adapted so researchers can read and listen to narrators' voices in terms of their own interests" (Bryne et al., 2009, p. 69). VCR coheres with a feminist orientation to knowledge production insofar as it emphasizes the relational and embodied nature of the words and actions of individuals. VCR is an appropriate method for our study because it aligns with many of the overall purposes of the course, which included an emphasis on the relevance of research positionality, power, and ethics in relation to data analysis and communication. Despite the promising nature of VCR, we were unable to engage in an in-depth use of the methodology due to time and resource constraints of the student researchers. Therefore, we limited the use of VCR to a semi-structured format wherein the instructional researchers posed the four stage process as guiding questions within asynchronous communications in order to scaffold data analysis among the student researchers.

Findings

Three organizing themes emerged from our process of engaging in participatory, abductive inquiry guided by the VCR method. The themes were: (1) data and data processes as stories, (2) data as an epistemologically and ontologically complex phenomena, and (3) collaborative data analysis as an alternative to data analysis as an individual, technical endeavor.

Data and Data Processes as Stories

Student statements about what they were learning about data resonated with the literature on sociopolitically attuned, critical approaches to data literacies/science education, which emphasizes the influence that people can have on processes of data generation, analysis, and communication. Their statements, however, tell a richer story than simply reiterating the fact that

there is always some degree of bias in data. Instead, by framing data and data processes in terms of stories, the students emphasized how people—influenced by their disciplinary backgrounds, personal experiences, and emotions—shape and become shaped by data and data processes.

Adele, a computer science major who was drawn to the storytelling dimension of the course, said,

> I feel like with every story now, there's a meta-story of who wrote it and how. That impacts the story. And to me, before data was just hard facts, but now there's so much more behind it that I didn't think about before. So it's almost like for every story that's told, there are stories outside of the story to think about.

Her statement suggests an opening not only to understand data processes as a way of telling stories about the world but also to understand data processes themselves as a form of story, where "there are stories outside of the story to think about." This point was emphasized throughout the course, first in introducing students to the idea of researcher positionality and then in engaging students with assignments that asked them to consider the role that people and their use of materials play in data processes. Readings such as D'Ignazio and Klein (2020) framed data processes in relation to broader discourses about data as well as subject matter-specific discourses about topics that data are connected to. The idea of data processes as a form of story was in fact initiated by the students themselves during group and individual interviews. This suggests that the course readings and assignments that sought to broaden students' conceptions of data processes were, at least in part, taken up by the students. Interestingly, although the instructional researchers did not ask the students to read literature such as Wilkerson and Laina (2018), Adele's statement along with the remainder of the statements in this findings section resonate with Wilkerson and Laina's tacit calls to promote a storytelling lens with data-intensive learning environments. Adele's statement that "there are stories outside of the story to think about" gestures towards Wilkerson and Laina's (2018) characterization of

data as a "messy, human endeavor" (p. 1224) and Shaughnessy and Pfannkuch's (2002) question posed to students, "What is the story contained in the data?" (p. 257). In each case, a storytelling lens encourages and provides a lens through which to see data as more than "just hard facts."

One aspect of data processes as a form of story involves the relevance of the disciplinary background of the person or people engaged in such processes. Bobble, a second year student majoring in public relations, stated,

> I think that in trying to represent data at all, you change it in the creation of that data. Well, you changed it because you got all sorts of bias leaking from all of your pores. And, like, I had to code it. In order to physically represent the data, I had to put it into categories. And that's, completely changing the thing, because [Saba] could have done it and she could have had completely different categories than me. And that would have entirely changed how our presentation works. And, I mean, I guess you can say that a representation of it is just in one interpretation of the data.

Here, Bobble refers to her work coding the survey-based data that she and her group members gathered on undergraduate students' opinions about COVID-19 masking in various public spaces. Bobble later explained that as she was helping to design the survey questions and, later, helping to analyze the survey responses, she considered her own background as a public relations major:

> If opinions are king, changing opinions are God. Our ability to measure a change in opinion is how we measure if PR works at all. A campaign is useless if it doesn't change opinions, whether that's growing awareness or incentivizing a buy. It's what we're trained to value. Surveys are a tool PR practitioners use to make sure we can show this. If we can't, we can't prove to our management that we're useful. Worth the money, worth the position, worth the time. So when I first sat down to consider how I was going to use the data from the survey, that is what I gravitated towards. How did their opinions change after being exposed to X? Most of the time, I do it for a social media post or blog, but a visualization isn't that far off, is it?

Bobble recognized that all people have different backgrounds and that bias is spread over everything that we touch. This is indicated by what we choose to omit, choose to include, choose

to edit, or choose to use. In Bobble's case, her background in public relations shaped her contribution to the design of her group's survey questions, which sought to understand how students' opinions on COVID-19 masking changed based on various settings and based on their exposure to a graph on COVID-19 cases in the county where they were located. Both of Bobble's statements suggest that if one had a different disciplinary background, then the outcome of survey design, data generation, data analysis, and data communication would have differed. The idea that her groupmate, Saba, may have coded the data differently had she taken the lead was particularly important to Bobble.

Constance, a psychology major who played a significant role designing a data visualization that was incorporated into her group's survey questionnaire that asked undergraduate students about their opinions on domestic and global food insecurity, added a related perspective when she stated:

> It does make me acknowledge just the sheer amount of ways that people can present their ideas and perspectives. It's been interesting seeing how difficult it is for one person to present data on their own in a widespread way. As I've gone through this course, I've realized the importance of having that graphic design background and there's just a lot that goes into presenting stuff in a way that people will really resonate with. People can come up with these super crazy, intricate things that are meaningful to them, but it might not be meaningful to the wider public, which is kind of scary in my perspective. It's just hard to balance that.

Despite the fact that Constance did not have formal graphic design experience, she gravitated toward the design aspects of visualizing data, noting the "super crazy, intricate" data visualization innovations to which the students were exposed through the data exploration and data postcard activities in the course. Her words suggest the relevance of a person's emerging disciplinary interests in how they approach the work of creating what are often considered neutral and objective representations of data.

Frodo, a student majoring in international relations within an interdisciplinary residential learning community focused on public affairs, recalled her involvement in a class discussion about an animated data visualization by Periscopic (2018) that depicted individual gun deaths in the United States and the years that were lost by each individual as a result. Frodo stated,

> In previous experiences, subjects like statistics and language arts were kept distinctly separate; if data points were mentioned, it was in something like history class which mentioned casualties, but even then it was acknowledged that statistics of death and war don't fully register with the human consciousness (one death is a tragedy, a million deaths is a statistic). This course, in essence, was all about that very occurrence; it aimed to orient ourselves with the data and realize that all data is reflective of lived experiences.

Here, Frodo points out that her previous experiences with data as merely numbers—a remark shared by nearly all of the students—were in part caused by the disciplinary boundaries imposed in schools that keep "subjects like statistics and language arts…distinctly separate." The interdisciplinary nature of the data storytelling course, on the other hand, encouraged student examination of innovative uses of data derived from a combination of disciplinary traditions. Periscopic (n.d.), for instance, characterizes itself as a "socially conscious data visualization firm" with expertise in design as well as experience in data analysis. Data processes, therefore, are shaped not only by people's disciplinary backgrounds but also by opportunities for *inter*disciplinary engagements with data, which highlight the varied sources of influence that can shape how data are generated, analyzed, and communicated. For Bobble, Constance, and Frodo, the variety of interdisciplinary perspectives, "reflective of lived experiences," aided in helping them understand the "stories outside of the stories to think about" with respect to data and data processes. This differs slightly from and builds on how the NASM (2018) report and other documents frame the interdisciplinary aspects of data literacies/science education. For NASM (2018), interdisciplinarity compels a "unified approach" to data literacies/science education,

where educators from varied backgrounds can work together to ensure that data-intensive learning environments do not become a collection of disjoint perspectives and skills. The statements from Bobble, Constance, and Frodo suggest a more asset-based perspective on interdisciplinarity. Heterogeneity of perspectives, held by students as co-authors of one another's learning, can be a driver for understanding data and data processes in richer ways. The fact that "people can come up with these super crazy, intricate things" can facilitate students' learning from one another, though Constance also suggests caution that one must also ensure that their perspectives about or representations of data are understandable to others.

Some students suggested that the ability to influence data processes is not inherently bad or something to be avoided. Saba, a premedical student enrolled in a science-focused interdisciplinary program, stated,

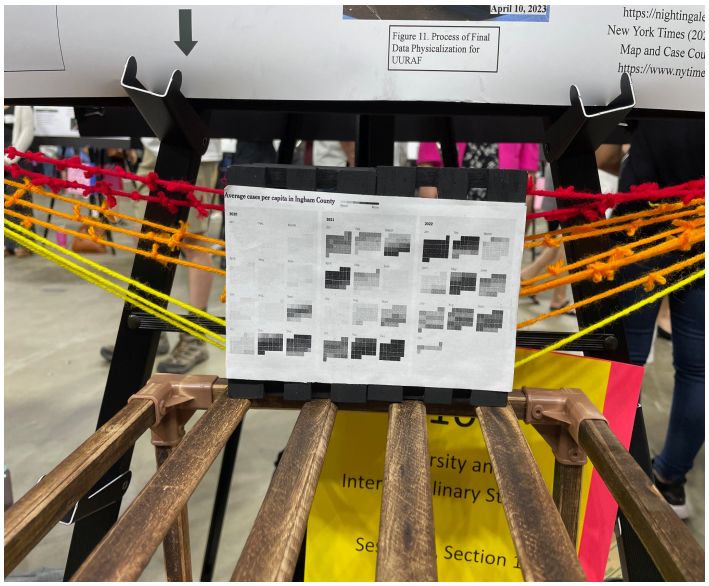
> This myriad of [data storytelling] approaches is important because it gives the person creating or collecting the data several different approaches to choose from, and it also allows them to truly express what their claims are and how the data enhances their story. Generally speaking, it allows them to create their story however they want. This is so important because without these approaches, all data stories will be like what I used to think data were - just numbers and graphs. Without context or a form of presentation, it will make data stories difficult to understand and when a story is hard to understand, it never really sticks with the audience and never produces an effect.

Saba is noting the affordances of data storytelling as a tool for making and expressing claims. Rather than framing this as a hindrance toward objectivity or worrying that it is an invitation that anything goes, she suggests that data storytelling is a way to help people think of data as more than "just numbers and graphs." She places an emphasis on the "effect" that data storytelling can have on audiences and on how one's approach to data storytelling can shape that effect. Once again, Saba's perspective resonates with Wilkerson and Laina (2018), who argue for the pedagogical benefits of drawing connections between storytelling and data processes.

Further, Saba's words indicate that data representations are more than the mere transmission of information. Instead, they are shaped by individuals and communities who must consider their effects on others. This idea is suggestive of perspectives within the literature on data literacies/science education that seek to challenge assumptions that data-savvy individuals must position themselves as detached, distant, and disembodied observers of phenomenon (e.g., D'Ignazio & Klein, 2020). Echoing Saba's sentiments, some students remarked on how their engagements with data were shaped by their personalities and emotions. We conclude this subsection by returning to Bobble and briefly discussing her remarks on the final project that she and her group members created for the data storytelling course.

Bobble, Saba, and one other student, Daisy, were part of the same three person research team for the survey-based projects. Together, they asked undergraduate students to share their opinions about COVID-19 masking in public spaces. The culmination of their project resulted in the creation of a data physicalization made out of yarn (Figure 11). Data physicalizations constitute "the practice of mapping data to physical form" (Bae et al., 2022, p. 1).

Figure 11. Yarn-Based Data Physicalization



Bobble, in discussing her role in the project, stated,

> I have anxiety. This came into play when it came to the yarn used in my physicalization. What we used was slippery– it didn't hold with glue, and tended to fall off the box used. But I didn't request different material, because the idea of it made me anxious. Which forced me to come up with a different way to anchor it, which ended up being knotting it to the box, which gave me the idea to add knots to the strands so the visually impaired could interact with it.

Emotion and bodily experiences of anxiety, therefore, played a role in the outcome of their research project, which was a fact that was not rejected by Bobble's instructors or peers for a lack of objectivity but rather embraced as part of the "meta-story" that constituted her group's data process. Indeed, as Frodo stated in a later writing, "Each study can be seen as a study of ourselves, in a certain sense." In a later correspondence, Bobble dug deeper into how her emotions played a role in her process of learning about data and how these emotions connected to broader sociopolitical issues.

> I was worried, quite often, about people's perceptions of it being 'childish' because it was made of yarn and wood and paper. It felt like a craft– and so I immediately discounted it in my own brain as being a worthy endeavor. It felt like I was silly, amongst all the other presentations, like the podcast. That shame is important to me, in how it was created. I finished it despite the knee-jerk shame of making something of yarn and wood, and so many times I almost dropped the idea for something more 'acceptable' and 'scientific' that fits the idea of data. That connects a lot with femininity and queer culture to me: changing oneself to fit the cultural frameworks you're expected to be in. Being more 'scientific' (whatever that means), despite your natural leanings towards yarn and paint.

Here, Bobble shares her insecurities about the particular yarn-based form of her and her group's final project, in part comparing what they produced to what one of the other three-person groups in the data storytelling course made, which was a podcast that we discuss in a later part of this paper's findings. Bobble's perspective aligns with and reiterates D'Ignazio and Klein's (2020) calls to embrace emotion and the body as essential and unavoidable aspects of data-driven processes. Moreover, rather than thinking of emotions in politically neutral terms, Bobble

connects her emotional reaction and the way she processed her emotions to issues of femininity and queerness. The data-driven learning process raised for her questions of shame and conformity to the disciplinary expectations of data-driven analysis, which are too often framed in terms of what is "acceptable" or "scientific." This broader sociopolitical connection resonates with D'Ignazio and Klein's (2020) argument that data and data processes are not merely a matter of emotions and the body but also about broader issues of the politics of knowledge production, where data-driven thinking is commonly associated with masculinist fantasies of disinterested objectivity and "hard facts" to the exclusion of other ways of knowing and being.

Data as an Epistemologically and Ontologically Complex Phenomena

The data storytelling course offered a unique opportunity to hold student discussions about the epistemological and ontological nature of data. These affordances arose as a result of framing data in terms of stories, provoking the students and instructors to pose questions about what data enable us to know or not to know (epistemology) or what data are in the first place (ontology). In one such discussion, Frodo stated,

> I began thinking about how even mundane data, such as my grocery habits, reveals much more than just my fulfillment of basic survival. My choices are influenced by personal preferences, childhood experience, targeted ads, cultural background, and socioeconomic status…Something as simple as feeding myself, which is literally the bare necessity of human experience, contains vast data implications and trends.

Frodo is suggesting how researchers cannot remain separated from points in data generation because of the interconnected nature of phenomena, with the result that despite our best efforts, we cannot help but influence our data and be influenced by the data that we and others generate. Daisy, a premedical student with prior experience participating in research projects connected to the fields of psychology and neuroscience, noted her initial and current impressions about data:

> At this point in the class, and partially now, I still thought of data as numbers. When I saw the qualitative data, I saw coding. The instinct was to mold what I had into numbers, so it could become data, although I knew that it was already data.

She then explained how interesting questions about data were raised for her through the data postcard activity that preceded the class's engagements with their survey-based research projects. In particular, Daisy used the data postcard activity to keep track of how often she checked her phone and for what purposes. Daisy stated,

> I never thought of it as data. But now that I have, something that we also talked about in classes, if we didn't track it before, is it still data? Which is an interesting question. I honestly don't have an answer to that right now…just because we don't collect it doesn't mean it's not out there. Um, and so that's interesting for me to think about.

Daisy raises a novel ontological question about data that are not necessarily central topics in other courses that aim to develop students' data acumen. Her question relates not only to the definition of the term "data" but also to what the boundaries of data are. The data postcard activity raised questions about what are and are not data, specifically in terms of their existence. Do data exist before we "collect" it? Daisy, in a later writing, asked,

> I noticed that some of my other classmates chose to do other activities, such as the amount of water they drank. Therefore, that data of theirs exists. The existence of my phone data or their water drinking data isn't really open to debate. However, does my water drinking data exist? Surely I drank water so data could have been collected about it. But I did not. Here is a gray area of the existence of data. And looking more broadly, does everything that has happened that could be collected (but wasn't) be considered data? Where do we draw that line of the existence of data?

Daisy suggests that the course left her with more questions than answers, going so far as to say in another writing that "data is quite honestly an undefined term, and the limit for what a data-driven story is, is very gray." Eve, a communication arts major with prior research

experience taking an Advanced Placement course in Research in high school, similarly raised an emerging uncertainty about what counts as data. She stated,

> My understanding of the term has changed because now I catch myself sitting here and I'm like, oh, data. I have to think about it, and I'm realizing I used to consider data a straight number that usually had a decimal point and was really specific. And then, you know how you talked about the clock? That sits in my head all the time. The clock example: How data can be an actual thing or data is subjective and all of that. So my understanding has changed, and I honestly right now could not define the term data because it lowkey still confuses me.

Like Daisy, Eve had originally conceived of data in terms of simple numbers. Through the course discussions and activities, however, her ideas about data changed, particularly through an example that Author 1 shared in class that used a household analog clock as an illustration of how something as simple as "10 + 4" can be said to be equivalent to "2" within an appropriate context. Likewise, one can only interpret a particular piece of data within a particular context. Her statement that data are "subjective" suggests an emerging understanding that data as a "straight number" does not sufficiently capture the complex processes from which data emerge and by which data can be analyzed, interpreted, and communicated.

Frodo, Daisy, and Eve's words echo much of this paper's previous findings, as well as the literature on data literacies/science education that emphasize the relevance of the backgrounds, experiences, and emotions of individuals and communities in relation to data and data processes. They build on such findings, however, by raising the following question: if data are not "straight numbers," then what exactly constitutes data? Instead of arriving at any definitive answer, Daisy and Eve's responses echo Mertala (2020) that the term data is broad and possesses no agreed upon meaning. Instead, following Rubin (2020), Frodo, Daisy, and Eve suggest that data are complex phenomena shaped by and shaping questions of when, where, why, and how data are generated, analyzed, communicated to others, and used in particular contexts.

These questions, among others, appear to govern the contours of what counts as data. Moreover, what counts as data gives rise to further questions, such as was the case when Daisy asked if "everything that has happened that could be collected (but wasn't) be considered data?" Posing such questions is indicative of the kind of abductive, storytelling approach to knowledge production that centers moments of surprise and provokes new questions for further investigation (Matei & Hunter, 2021).

Related to these epistemological and ontological questions about data is the issue of power and ethics in data. Many students acknowledged that the results and interpretation of data can be influenced by factors such as one's disciplinary background, lived experiences, or emotions, thereby raising questions about what we can and cannot know from data and about what data are. Some students connected these questions to broader issues of power and ethics. Sasha, a student majoring in English and psychology who talked about her interest in eventually working in the field of psychiatry, stated,

> I learned about what's behind data and who is behind data, which I think is even more important if you think about that because who is behind data is always going to have the most control and the most power, because they are the people putting it out there, deciding how it's going to be put out there and how people are going to perceive it, decide what it's going to be…and that is very important.

It is not merely the case that data are influenced by people. Sasha points to one consequence of such influence, namely that those responsible for processes of data generation, analysis, and communication can make decisions that influence others. Data may be unique in this sense, insofar as its connection to mathematics and statistics can create the impression of authority and therefore lend weight to claims made with respect to such data (Rubel et al., 2021).

Sasha continued to remark on the connection between data and power by drawing on her positionality as a biracial Middle Eastern and white woman. Referring to the data exploration

assignment where students were asked to identify and examine news articles that used data to tell a story, Sasha stated,

> I had a very hard time finding a lot of data stories based off of women, especially minority women, which I feel like unfortunately is just like a given now, and it really shouldn't be. But a lot of the data studies, they're going to be done on men. And if they are going to be done on women, typically it's harder to find them done on minorities or specific minorities. I personally am able to find, you know, a good amount of data stories on maybe just women in general, or Black women even. But it's harder for me to find anything about Middle Eastern Americans or Arab American women especially. I can't find really almost anything…It's hard to say, hey, this is what happens in the medical field all around for everyone. That's not how that works. Women are just, we're built different, we're a different breed. So we need a different type of data and a different data set because it's not the same. It just doesn't work like that.

Likewise, Adele stated,

> Another data story which affected my positionality was one that was brought up in class about women athletes and how they're often viewed as just small men. I didn't realize that being a woman was such a disadvantage when it came to data collection because it appears that a man is the default in data collection and there's so much more data about them. This can cause problems, especially in sports medicine since women need to be treated differently than men since their bodies are different and viewing them as little men can be super damaging and cause them to get improper treatment that isn't right for their bodies. I still view my salient identity markers much as the same as the beginning of the course. However, I now see what impact they have in data collection since many groups such as women or minorities are often underrepresented. So some of my identities, such as my race, bring me power in the world of data while my gender does not.

Sasha and Adele are referring to an important topic that was only briefly raised in class and deserved more time and attention: the relative scarcity of data-driven studies that take into account the unique circumstances of women and racialized minorities, such in the case of studies that were brought up in class about sports medicine or car safety. The connection between power and data, therefore, did not merely refer to power as a form of influence but also to systemic forms of power that manifest in patterns of discrimination, hierarchy, and oppression cutting across lines of race, gender, ability, sexual orientation, socioeconomic status, and other identity

markers. Sasha and Adele's discussions resonate with critical approaches to data literacies/science education that call for opportunities for students to draw connections between data and "deeper structural issues of inequality" (Fotopoulou, 2021, p. 1641). The students in this study's data storytelling course were able to use activities such as the Dear Data project and their year-long survey-based projects to understand data as a sociocultural process and see themselves as agentic producers, rather than passive collectors and consumers, of data (Hardy et al., 2020; Stornaiuolo, 2020). Notably, while studies such as Stornaiuolo (2020) lamented the fact that student identity development appeared to take precedence over deeper interrogations of structural inequity within the data-intensive learning environment, Sasha and Adele's remarks—along with those of Bobble mentioned above—suggest that the students gained some inroads into developing greater sociopolitical consciousness with respect to data and data processes than they had before. Many factors could account for the difference in these findings as compared to those in Stornaiuolo (2020): the slightly older age of the students in this study as compared to the high school-aged youth in Stornaiuolo (2020), the exposure to explicitly political approaches to data representation such as those by Chalabi (2020) and D'Ignazio and Klein (2020), or simply the types of questions that the instructional researchers and students posed or the topics they raised during class discussions, focal group interviews, and individual interviews. Despite these findings, the course design could have placed an even greater emphasis on critical perspectives with respect to data and data processes, as we will discuss in this paper's concluding remarks.

Collaborative Data Analysis as an Alternative to Data Analysis as an Individual, Technical Endeavor

The students' year-long survey-based research projects entailed creating a data story that, in most instances, involved active collaboration among the students, both in terms of analyzing

their data and creating a data-driven story based on the results of their analysis. Many students noted their experiences with and affordances of engaging in data analysis in collaboration with their peers, as opposed to merely applying statistical techniques to data in isolation. Dawn, a business major who was formerly a part of the same interdisciplinary program in which Frodo was enrolled, discussed her group's process of creating the survey questions for their chosen social issue, which sought to understand undergraduate students' opinions on the topic of economic mobility in the United States. Dawn stated,

> In the process of crafting research questions, my group compared our work to other teams within the course. I found myself asking questions such as: Why did I not think to ask that question? Why would they prioritize asking that question when there is limited space? Are we all missing a question that would enhance everyone's findings? These questions made me realize that the unique individualism of each member of the course pushed me to question the credibility in my interpretation and formation of data points. Even the questions I was asking myself were biased because they were the only questions I thought to complete due to my personal experiences. Communicating with a diverse group of researchers and peers led me to understand that the origin of the process that produced data questions is almost if not as important as the data result itself.

The comparison that Dawn refers to was made possible by an important constraint that emerged in the early months of the course. In particular, the student and instructional researchers were aware that we would need to receive approval for disseminating our survey not only from our university's institutional review board but also the university's deans. Therefore, although our class formed three groups of three with the intent to conduct three separate survey projects, we decided to consolidate all three projects into a single survey, with each page of the survey containing questions that aligned with one of three social issue topics proposed by each group. Each group, then, was able to see one another's questions and, through various class structures, collaborate with other group members to brainstorm and build off one another's questions. Dawn's statement that the "process that produced data questions is almost if not as important as

the data result itself" aligns with and specifies the NASM (2018) report recommendation to approach data literacies/science as a "general approach to problem solving" (p. 20). Moreover, following Dawn's words, we argue that question generation—particularly within collaborative settings that encourage students to ask "Why did I not think of that question?"—is an essential skill that should be added to the list of data acumen skills (NASM, 2018) and dimensions of working with data (Rubin, 2020) that scholars promote for data literacies/science education.

The theme of collaboration appeared again while Dawn and her group members met to develop their data story based on their survey results. They had initially planned to create a TED talk in order to tell a story about undergraduate students' opinions on economic mobility in the United States. They later changed their data story to be in the form of a podcast. Eve, who was part of Dawn's research team, stated,

> I would say the podcast was, I think, fitting for our type of research. I think it did a good job of, it helped us organize our thoughts for the presentation. So not even as a finished product, but as a means to an end, I think that was probably the most helpful in seeing where all three of us thought the same and thought differently about certain things. [Dawn], obviously she was originally, she was a [student in the same residential college as Frodo]. I'm not sure what her major was, but she had thoughts on income inequality and honestly, I didn't even fully understand it, but she and [Adele] had different thoughts on the intro and on one of the questions. I don't even remember where I stood because it was helpful to kind of get organized and get our thoughts together. And also, I think listening to it back, you could learn a lot, even as someone doing the research, but I just don't think it was the most productive. I think it could have been used as a thing to get organized, but it wasn't the most productive final project of displaying our information.

The collaborative setting of co-creating a podcast, in other words, was a means of developing greater familiarity with their data as a result of shared inquiry. Such shared inquiry, consistent with this paper's previous finding, afforded the opportunity for conversation across disciplinary backgrounds, as Dawn, Adele, and Eve all had different academic interests and future career goals in mind. Interestingly, Eve notes that the podcast, as a final product, was not

"the most productive." In a separate statement, Eve explained that this was because when they presented their work at the undergraduate research conference, the judges and other attendees who visited their poster only spent a few minutes listening to the podcast. Therefore, from Eve's perspective, it was not the most public-facing and engaging form of data-driven story that they could have chosen to use for purposes of the research conference setting. The moments of shared inquiry that Eve alludes to builds on findings suggested in Wilkerson and Laina (2018), where students worked in groups to analyze and represent geospatial data of their city. In particular, whereas studies such as Wilkerson and Laina (2018) merely provide brief references to collaborative data analysis, Dawn and Eve's words paint a more specific picture of what collaborative data analysis entails, what thoughts the students considered as they worked together, what affordances they identified, and what limitations arose.

Adele elaborated on her group's experience creating a data-driven podcast together when she stated,

> A lot of things came up in our podcast that we'd never even really discussed as a group before because that's just where the conversation led. And so I feel like kind of inserting my opinion and my background and what I've kind of learned through the research process. I think it was very important using my voice in order to kind of move it along because then that gave something for Eve to comment on or about her experience. And same with Dawn. So I think the three of us, I think we did a very good job of all using our voices and all kind of sharing an equal amount and sharing our backgrounds and our opinions. Because that gave room for someone else to say, well I agree with you, well I disagree with you a little bit. Or well, like, I see where you're coming from, but, and so I feel like using our voices and our backgrounds was super important in keeping the conversation going so naturally and kind of even leading to things that we hadn't really even discussed as a group yet.

Adele's use of the word "voice" highlights the importance of collaboration as a tool to help move data analysis forward. Each of Adele, Dawn, and Eve's voices consisted of agreements, disagreements, and differences in perspectives that constituted an ongoing process

of sense-making with data. As part of their survey, the group had asked undergraduate students to read a graph on economic mobility and opine on whether the "American dream" was more or less attainable than in the past. The group had assumed that students would believe that the American dream was "significantly more attainable" because of technological advances in society. Dawn, in later writing, stated,

> This assumption was squashed when the data revealed that 51.4% of respondents believed that the American dream is significantly less attainable than it was in the past. The group chose to put this at the first data point on our poster because it represented the power data can play in communicating the influence the difference our means of socialization and positionalities play into our interpretations and assumptions. This drove my team to dig further into the reasoning behind why we had made our assumption so quickly and confidently and found that we shared many common factors in our positionalities that pertained to economic mobility: we all came from upper-middle class families, are white families, and are currently pursuing an undergraduate education. The takeaway was so important to the team as individuals and researchers that we chose to make this our first talking point on the podcast agenda. It interests me how the influence of socialization differs in the various residential colleges we interviewed, and I wish we were able to know which college the respondents identified with. Unfortunately, this was a limitation of the study.

The role that communication plays is critical in understanding and conveying the role that influences such as researcher positionality plays across all stages of the data-driven process. Without diverse voices present in each of these stages, as Dawn suggests, we cannot understand what we do not know. Bias, in other words, is not simply inserting persuasive language into a survey question or into one's interpretation of data. It is also the unintentional communication of one's socialized experiences that occurs when one speaks or writes from their individual perspective. The conclusions that Dawn, Adele, and Eve's group made would not have been possible without communication among the team members, course members, and survey respondents. Such communication was necessary in order to tackle the critical and creative task of designing any data-driven story, such as a podcast or a data physicalization made out of yarn.

These reflections by Adele and Dawn align with and build on the literature, not only insofar as they reflect an understanding of data as an indicator of aggregate trends (Konold et al., 2015) but also that they reveal how student thinking with respect to data need not be seen as an individually held trait. Instead, how students think about data can and should be considered to be a co-produced phenomenon shaped by specific patterns of interaction situated within students' past experiences, future goals, and broader discourses about data. Discussions of data analysis as a co-produced phenomenon, however, are not prevalent within the literature on data literacies/science education and therefore is a topic that merits further consideration.

Discussion and Conclusion

In this paper, we address the question of how opportunities for critical, creative, and interdisciplinary engagements with data shaped how students made sense of data and data processes. The idea that data processes are shaped by influences outside a dataset itself is a major theme that emerged throughout the interviews and written assignments by the students, as well as the participatory data analysis that forms the methodological foundation of this paper. Such finding supports recommendations found in policy documents such as the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report* (GAISE College Report ASA Revision Committee, 2016), which states that "[S]tudents should become *critical consumers* of statistically-based results reported in popular media, recognizing whether reported results reasonably follow from the study and analysis conducted" (p. 8, emphasis in original). Moreover, statements from students such as Frodo that "data is reflective of lived experiences," as informed by their engagements with dynamic data visualizations such as Periscopic's graph on gun deaths in the United States, suggests that students successfully moved between seeing

data as a pointer to specific events, a value for individual cases, and a vehicle for communicating frequency and aggregate information (Konold et al., 2015).

Students developed an emerging understanding that data are not "straight numbers." Some students elaborated on philosophical discussions about data that were raised during class discussions, such as the question of whether data exist before they are measured by people. Some students expressed confusion about the boundaries of what counts or does not count as data or a data-driven story. Other students noted how data processes can confer power. Moreover, the lack of availability of certain types of data, especially data that could be used to better understand the needs and experiences of women and people from racially minoritized backgrounds, highlights how power is always at play in data processes despite perceptions that data speak for themselves. These findings align with research that emphasizes the need for a critical approach to data literacies/science education that challenges naive assumptions about the objectivity and neutrality of data (e.g., D'Ignazio & Klein, 2020; Fotopoulou, 2021). Moreover, these findings go beyond literature such as Stornaiuolo (2020) by providing some brief insight into students' emergent sociopolitical consciousness around the connection among data, data processes, ethics, and power.

Lastly, students noted the importance of collaboration. Collaboration and communication became important aspects of the course early on when students shared their data postcards with one another and provided feedback for revisions. As Dawn, Eve, and Adele stated, collaboration became a defining feature of their data-driven stories, where communication and the shared use of their voices supported them in making sense of their data, even more so than creating a finished product for their research projects. These findings, marked by the specificity with which Dawn, Eve, and Adele described their collaborative analytical engagements with one another,

suggest the need for further research to fill in a gap in the literature on data literacies/science education, which often frames skills and dispositions with respect to data and data processes as individual traits rather than collective accomplishments.

These findings also suggest that the interdisciplinary nature of the course—along with its emphasis on critical and creative approaches to data generation, analysis, and communication—presented unique affordances and challenges for student learning. Because the course did not have prerequisites in mathematics, statistics, or computer science, students entered the course with a wide range of backgrounds, beliefs, skills sets, and career aspirations. This diverse coming together of backgrounds, coupled with class readings and discussions that made the relevance of researcher positionality explicit within data processes, created conditions for students to understand that data do not exist in a vacuum but are instead shaped by broader forces that are responsible for how data come into being and become transformed into substantive claims about phenomena. These conditions of interdisciplinarity also enabled the students to choose salient social issues that resonated with their interests and their undergraduate student peers, thereby creating an authentic context for data-driven inquiry that had the effect of maintaining student engagement. This resonates with the NASM (2018) report's call for a unified approach to data literacies/science education that leverages interdisciplinary expertise to promote data literacies/science as an approach to problem solving rather than a disjoint set of skills. Whereas the NASM (2018) report, however, suggests that educators from interdisciplinary backgrounds should be the drivers of such unified approach, findings in this study appear to "flip the script" and promote a view of interdisciplinarity grounded in *students'* varied backgrounds rather than those of any particular instructional community.

Despite these affordances of the course, the students and instructional researchers also faced challenges as a result of the course design, along with other external factors. The interdisciplinary nature of the course was a strength but also a potential impediment to teaching and learning because it resulted in course goals that were arguably too broad or lacking in clarity. Data acumen encompasses multiple skills and dispositions, ranging from data generation and acquisition, management and curation, modeling, representation, communication and teamwork, exercise of good judgment, and responsible and ethical use of data (NASM, 2018), along with a critical awareness of the limitations of data-driven analysis and an understanding of the historical and contemporary harms caused by data-driven processes (e.g., Fotopoulou, 2021; Stornaiuolo, 2020). The data postcard activity and survey-based research project offered opportunities for students to develop and engage in all of these aspects of data acumen. Coupled with the fact that class discussions were often student-led, and the instructional researchers repeatedly adjusted instruction according to students' developing understandings about data, it was unclear at times what students were expected to focus on at any given period of the course. How much should the students have focused on using statistical techniques while making sense of their data postcards? Or was the focus of the activity on learning to collect, manage, and visualize data? What particular skills and dispositions about data visualization should be emphasized if students were free to create and experiment with any form of data visualization?

It should also be noted that the data storytelling course's emphasis on criticality and creativity were not universally taken up by all students in the course and at all times. Eve, for instance, suggested that her experience in the class did not align with her initial expectations for a course about data:

> Coming into this class and taking a different approach was off putting. Maybe it would have been better if I didn't take AP Research because I think I had a fixed

> mindset on what my research was supposed to look like…I didn't like the postcards, and that was maybe just me. I think I had other things going on, but I didn't know how to creatively display my work at all. And those kids came into class and they had these really pretty cards with all this information on it, and it looked good, and I was like, I don't see how you managed to take these numbers and turn it into this. So I guess that was just the lack of creativity that got me there.

Her statement challenges assumptions that creating opportunities for creative engagements with data will necessarily create or maintain student interest. It could be the case that the notion of creativity in the course was too narrowly defined to encompass the use of the arts and aesthetic novelty within data-intensive contexts. Broader notions of creativity might have created more engaging learning opportunities for students such as Eve, who in the same interview discussed her preference for data stories presented in the form of research reports. Thus, "creative" approaches to data generation, analysis, and communication could include thoughtful approaches to working with data in traditional contexts, such as incorporating new and strategic ways of combining issues of power, ethics, and researcher positionality within the research report format.

Lastly, the demographic characteristics of the class may have served as both a strength and limitation for student learning in the course. The course counted as an honors credit through the university's Honors College, thereby limiting the types of students who would enroll in the course. Moreover, the demographics of the enrolled students reflected the demographic composition of the university, which is a predominantly white institution. This may have had the effect of narrowing the availability of diverse perspectives in the course, which in turn likely shaped the type of discussions that were raised in class, student responses to course readings and assignments, the social issue topics that were proposed for students' survey-based projects, the survey questions that the class designed together, and the ways the survey data were interpreted

within each research team. Dawn, Adele, and Eve suggest that this limitation was in fact present within their own group, as they acknowledged that their shared backgrounds may have shaped how they approached the topic of economic mobility in the United States. Further, the entire class was composed of students who identified as women. It is unclear why this may have occurred, and the role that students' genders played in the course is a potential topic for further study. One might argue that this factor created space to emphasize certain topics such as the lack of availability of health and car safety related data about women and, in particular, women from racially marginalized backgrounds. It might also be true that this aspect of the course led to or was caused by the course's embrace of researcher introspection, vulnerability, communication, and connection with respect to data literacies/science, a topic that has been associated with "masculinist, totalizing fantasies of world domination as enacted through data capture and analysis" (D'Ignazio & Klein, 2020). Nonetheless, it may also have been the case that having more students who do not identify as women would have contributed to more diverse perspectives.

REFERENCES

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D.A. (2020). *PreK-12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education*. American Statistical Association and National Council of Teachers of Mathematics.

Bennett, J. (2010). *Vibrant matters: A political ecology of things*. Duke University Press.

Bhargava, R., Deahl, E., Letouzé, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). Beyond data literacy: reinventing community engagement and empowerment in the age of data. Data-pop alliance white paper series. *Data-Pop Alliance White Paper Series*. Accessed at: https://datapopalliance.org/wp-content/uploads/2015/11/Beyond-Data-Literacy-2015.pdf

Bourke, L. (2009). Reflections on doing participatory research in health: Participation, method and power. *International Journal of Social Research Methodology, 12*(5), 457-474.

Brown, L., & Gilligan, C. (1992). *Meeting at the crossroads: Women's psychology and girls' development*. Ballantine.

Brinkmann, S. (2014). Doing without data. *Qualitative Inquiry, 20*(6), 720-725.

Byrne, A., Canavan, J., & Millar, M. (2009). Participatory research and the voice-centered relational method of data analysis: Is it worth it? *International Journal of Social Research Methodology, 12*(1), 67-77.

Cairo, A. (2016). The truthful art: Data, charts, and maps for communication. New Riders.

Cobb, G., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly, 104*(9), 801-823.

D'Ignazio C., & Klein, L. F. (2020). *Data feminism*. The MIT Press.

Draper, E. B. (2020, December 16). Data science job growth in 2021 and beyond. *The Commons*. Accessed at: https://lambdaschool.com/the-commons/data-science-job-growth-in-2021-and-beyond

Feigenbaum, A., & Alamalhodaei, A. (2020). The data storytelling workbook. Routledge.

Fotopoulou, A. (2021). Conceptualising critical data literacies for civil society organisations: Agency, care, and social responsibility. *Information, Communication & Society, 24*(11), 1640-1657.

Green, K. L. (2020). Radical imagination and "otherwise possibilities" in qualitative research. *International Journal of Qualitative Studies in Education*, *33*(1), 115-127.

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, *14*(3), 575-599.

Hardy, L., Dixon, C., & Hsi, S. (2020). From data collectors to data producers: Shifting students' relationship to data. *Journal of the Learning Sciences, 29*(1), 104-126.

Holland, S., Renold, E., Ross, N., & Hillman, A. (2008). Rights, 'right on' or the right thing to do? A critical exploration of young people's engagement in participative social work research. NCRM Working Paper Series 07/08. Accessed at: http://eprints.ncrm.ac.uk/460/

Ihde, D. (1979). *Technics and Praxis. Boston studies in the philosophy of science (Vol. 24)*. Springer.

Jackson, A. Y, & Mazze, L. A. (2008). Experience and "I" in autoethnography. *International Review of Qualitative Research, 1*(3), 299-318.

Kroes, P. (2003). Physics, experiments, and the concept of nature. In H. Radder (Ed.), *The Philosophy of scientific experimentation* (pp. 68-86). University of Pittsburgh Press.

Konold, C., Higgins, T., Russell, S. J., Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics, 88*, 305-325.

Lee, V. R., & Dubovi, I. (2020). At home with data: Family engagements with data involved in type 1 diabetes management. *Journal of the Learning Sciences, 29*, 11-31.

Lee, V. R., Wilkerson, M. H., & Lanouette, K. (2021). A call for a humanistic stance toward K–12 data science education. *Educational Researcher, 50*(9), 664-672.

Letvak, S. (2003). The experiences of being an older perioperative nurse. *Association of Operating Room Nurses (AORN) Journal, 78*, 635-647.

Loukissas, Y. A. (2019). *All data are local: Thinking critically in a data-driven society*. MIT Press.

Lupi, G., & Posavec, S. (2016). *Dear data*. Princeton Architectural Press.

Lupi, G., & Posavec, S. (2018). *Observe, collect, draw! A visual journal*. Princeton Architectural Press.

Matei, S. A., & Hunter, L. (2021). Data storytelling is not storytelling with data: A framework for storytelling in science communication and data journalism. *The Information Society, 37*(5), 312-322.

Mertala, P. (2020). Data (il)literacy education as a hidden curriculum of the datafication of education. *Journal of Media Literacy Education, 12*(3), 30-42.

Messy Data Coalition (2020). Statement. Accessed at: https://messydata.org/statement.pdf

National Academies of Sciences, Engineering, and Medicine (NASM) (2018). *Data science for undergraduates: Opportunities and options*. The National Academies Press.

Nind, M. (2011). Participatory data analysis: A step too far? *Qualitative Research, 11*(4), 349-363.

Owens, T. (2011). Defining data for humanists: Text, artifact, information or evidence. *Journal of Digital Humanities, 1*(1), 6-8.

Paliadelis, P., & Cruickshank, M. (2008). Using a voice-centered relational method of data analysis in a feminist study exploring the working world of nursing unit managers. *Qualitative Health Research, 18*(10), 1444-1453.

Pinto, K. (2004). Intersections of gender and age in health care: Adapting autonomy and confidentiality for the adolescent girl. *Qualitative Health Research, 14*, 78-99.

Quigley, A. (2020). Personal data postcard. *Introduction to Data Storytelling*. Accessed at: https://sites.google.com/macalester.edu/introtodatastorytelling/assignments/personal-data-postcard

Raffaghelli, J. E. (2020). Is data literacy a catalyst of social justice? A response from nine data literacy initiatives in higher education. *Education Sciences*, *10*(9), 233.

Rakotondravony, N. (2019, November 1). Power to the powerless: An interview with Mona Chalabi. *Medium*. Accessed at: https://medium.com/nightingale/power-to-the-powerless-an-interview-with-mona-chalabi-39d73647d80a

Rubel, L., Nicol, C., & Chronaki, A. (2021). A critical mathematics perspective on reading data visualizations: Reimagining through reformatting, reframing, and renarrating. *Educational Studies in Mathematics, 108*, 249-268.

Rubin, A. (2020). Learning to reason with data: How did we get here and what do we know? *Journal of the Learning Sciences, 29*(1), 154-164.

Sanches, P., Howell, N., Tsaknaki, V., & Helms, K. (2022). Diffraction-in-action: Designerly explorations of agential realism through lived data. *CHI Conference on Human Factors in Computing Systems (April 29-May 5, 2022)*. ACM. New York, NY.

Shaughnessy, J. M., & Pfannkuch, M. (2022). How faithful is old faithful? Statistical thinking: A story of variation and prediction. *Mathematics Teacher, 95*(4), 252-259.

Staunæs, D., & Brøgger, K. (2020). In the mood of data and measurements: Experiments as affirmative critique, or how to curate academic value with care. Feminist Theory, 21(4), 429-445.

Stornaiuolo, A. (2020). Authoring data stories in a media makerspace: Adolescents developing critical data literacies. *Journal of the Learning Sciences, 29*(1), 81-103.

Van Wart, S., Lanouette, K., & Parikh, T. S. (2020). Scripts and counterscripts in community-based data science: Participatory digital mapping and the pursuit of a third space. *Journal of the Learning Sciences, 29*(1), 127-153.

Watson, S. (2015). Data is the new '__'. *dis magazine*. Accessed at: http://dismagazine.com/discussion/73298/sara-m-watson-metaphors-of-big-data/

Wilkerson, M. H., & Laina, V. (2018). Middle school students' reasoning about data and context through storytelling with repurposed local data. ZDM, 50(7), 1223-1235.

Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, *12*(3).

CHAPTER 5: CONCLUSION

Being able to do my own visualization and being able to do a little drawing and color things and make it look fun and have something creative. I mean, there's a million ways that you could represent data. I would think of it more of a square thing, and now I think of it more as like a circular thing, if that makes sense. Just, it's like more fluid than it was before. Like it's a lot less rigid. - Constance

I begin this concluding chapter with the words of Constance, whose quote I included in the epigraph of this dissertation's introduction. Her words resonate with me because they evoke a desire that motivated how I approached this study as well as much of my graduate education, namely a desire to understand data "more as like a circular thing" rather than "more of a square thing." By engaging in this study, I sought to understand data in more fluid ways and to attune to how that opened up or foreclosed possibilities for teaching and learning within data-intensive environments. Her words also bring me back to my middle school classroom in Queens (Figure 12).

Figure 12. My Classroom in Queens

It was there that I began to understand mathematics as "more like a circular thing" rather than "more like a square thing," as more than a set of disconnected rules and rather something

with a life of its own, as an "Other[] in relation…rather than merely a context in which teacher-student relations occur" (Chen, 2023, p. 285). During graduate school, I continued to develop my emerging philosophical orientation toward mathematics, whether it was through coursework in Mathematical Ways of Knowing or seeking to understand mathematics from various relational ontological perspectives (e.g., de Freitas & Sinclair, 2013). My path moved toward an interest in data as a result of a multi-year research collaboration developing theoretical frameworks and commentaries in data science education (Kahn et al., 2022; Lim et al., 2022; Rubel et al., 2021), my data literacies/science work in the Interdisciplinary Inquiry and Teaching Fellowship, and my pursuit of a master's in Statistics during my doctoral program. These experiences led me to shift my interests toward data literacies/science education, and yet the same pull to understand mathematics in richer and more relational ways continued to inform my desire and approach of reimagining teaching and learning about data. In the remainder of this chapter, I provide a brief summary of the three central chapters of this dissertation, its contributions to and resonances with the relevant literature, its personal and broader impacts, and the questions and next steps it raises for me as an educator and emerging scholar.

Three Analytic Cuts

In this dissertation study, I engaged in three distinct analytic cuts to think about/with a relational ontological orientation to data and data literacies/science education. This section provides a brief summary of this dissertation's core chapters, each within which (my co-authors and) I sought to attune to and consider the implications of such a relational ontological orientation for data literacies/science education.

Chapter 2 attends to the ontological and epistemological orientations to data and data practices found within prominent data science education reform efforts. Drawing on the notion of

scripts and valences, I briefly identify three such orientations involving the themes of workforce preparation, risk management, and access and opportunity for all students. These themes are not exhaustive, and I raise them in order to provide context for provoking new ideas and questions for data science education based on a diffractive reading of agential realism, theories about noise, and the literature on data science and data science education. In particular, this diffractive reading enables me to reconsider how the material world and other "noisy" and liminal agents intra-rupt and become intra-rupted by the disciplinary concerns of data science and data science education. Central to my reading is the understanding that data, data practices, and data-intensive learning environments are highly relational. Noise is a broad and complex concept that imbues this notion of relationality with distortion, messiness, and ambiguity. In turn, noise, as read through agential realism, calls into question widespread preoccupations with using data to elevate signal over noise. In the alternative, I advocate for understanding data science and data science education as the intra-active entanglement of signal and noise together.

My analysis re-reads various existing analyses and initiatives related to data science and data science education. Literature that encourages teachers to create opportunities for students to experience material resistances with respect to data (e.g., Hardy et al., 2020) is an important step toward acknowledging the agency of the material world within data-intensive learning environments. This chapter's diffractive reading extends these ideas beyond matter as a mere mediating agent and toward matter as an agentic force within parasitic relations of data generation, analysis, and communication. I do not discount the value, however, of traditional data analytic techniques that have historically treated matter as a source of noise to be erased or controlled. I cite the example lessons in Weiland and Williams (2023) as an example of the need to strike a balance between elevations of signal and acknowledgments of noise. Noise, in other

words, is not necessarily positive or negative but rather, as Thompson (2017) states, it is a force relation that induces change. Noise, furthermore, inspires a reading of data science and data science education in terms of intra-ruptions. This notion emphasizes that disruption has no inherent meaning but is instead a notion co-constituted within phenomena. I use intra-ruption to nuance data science oaths of the kind promoted by the National Academies of Sciences, Engineering, and Medicine (NASM, 2018) that endorse a vision of ethics that equate control, candor, and compassion with social benefit, without considering how people from traditionally marginalized backgrounds are too often sidelined, forgotten, or oppressed by data-driven systems. They are, in other words, cast aside as noise for the sake of signal and clarity. Crawley's (2017) notion of Blackpentacostalism provides a generative lens with which to understand the potential of a joyful, liberative, and embodied conceptualization of noise in data science and data science education.

Chapter 3 shares vignettes of two students' engagements with Lego-based and yarn-based data physicalizations as part of a data postcard activity and their year-long research projects involving creating data stories based on the results of a widely disseminated survey. The vignettes highlight the materiality of data by drawing attention to how students' words and gestures, as part of the same phenomena as the agentic capacities of the more-than-human world, co-produced, at times, traditional onto-epistemologies of data and, at other times, new orientations toward data within data-intensive learning environments. The vignettes also highlight data's temporal dimensions, suggesting that data are always imbricated with broader discourses about time, uncertainty, and risk. In particular, moments emerged when the hands, gestures, materials, and discourses of the two students—Bobble and Saba—along with my own assignments and involvement in the course, made possible the idea that data was a pointer to an

externally existing reality such that data could be used to understand the world from a distant and detached perspective. These same moments, however, could also be read as suggesting non-representational possibilities for engaging with data. This was the case when strong affective responses emerged in Bobble and Saba's discussion of their data physicalizations; when Saba incorporated a cookie into her Lego-based structure; and when Bobble began a new data-driven story as she was disassembling her data physicalization. In such moments, Saba, Bobble, and the more-than-human material world enacted a participatory reworking of data that pulled data away from being an abstract pointer to an independently existing reality and toward a non-representational orientation of data as less discrete and quantifiable.

My analysis is grounded in a relational ontology of data and data literacies/science education. Consistent with my analysis in Chapter 2, a relational ontology grounded in agential realism (Barad, 2007) and Ingold's (2011) concept of meshworks seeks to go beyond a conception of classroom materiality as mere passive objects ripe for human manipulation and control. Instead, as an alternative to the cognitivist and behaviorist accounts of teaching and learning prevalent within the study of instructional design and pedagogy, a relational ontology emphasizes the co-constitution of all participants within a learning environment. In other words, classrooms are dynamic and contingent phenomena, in the sense discussed by Barad (2007). Following Bozalek and Zembylas's (2017) notion of response-able pedagogy, this chapter decenters approaches to instructional design and pedagogy that emphasize interventions based on evidence-based results. Instead, instructional design and pedagogy are a matter of taking up one's response-ability as a participant in a design process where all agencies are enacting boundaries and exclusions. One concrete implication for the everyday work of teaching and learning is to attend to and invite conversations about the normativity of disciplinary knowledge

and the reasons why one would want to attend to questions of epistemology and ontology in the first place.

Chapter 4 addresses the following research question: How did opportunities for critical, creative, and interdisciplinary engagements with data throughout the data storytelling course shape how students made sense of data and processes of data generation, analysis, and communication? The chapter focuses on three aspects of the interdisciplinary data storytelling course: (1) interdisciplinarity, (2) data as a tool to tell stories, and (3) data as a story. The chapter's standout feature lies in its methodology. The study combines elements of participatory data analysis and writing, abductive analysis, and a voice-centered relational approach to reading student work and interview data. This methodology enabled five of the nine students from the course—i.e., those who agreed to participate as co-authors of the chapter—to contribute to and share responses for this chapter's research question. Large group and individual meetings were held to solicit student ideas. The students identified particular transcripts and student work to highlight, discussed how they made sense of their own learning from the course, and contributed writing that appeared throughout this chapter. Because this process occurred approximately one year after the conclusion of the course, there was an arguably longitudinal aspect to their contributions.

Several organizing themes emerge from this shared analysis. These themes involve how the students understood data as a phenomena shaped by people and materials, what exactly constitutes the boundaries of data and data storytelling, how data is entangled with broader patterns of power and oppression, and how data processes can be understood as a collaborative endeavor rather than an individual activity solely focused on the application of computational and statistical techniques. The interdisciplinary nature of the course, coupled with opportunities

to engage with data in critical and creative ways, appeared to move students from seeing data as mere numbers on a spreadsheet and toward seeing data as a more complex and, in some cases, confusing concept. Part of this interdisciplinarity involved enabling students to engage in processes of data generation, analysis, and communication in an authentic context based on social issues that resonated with their interests and backgrounds. Despite the affordances of interdisciplinarity, it was also a constraint because it resulted in course goals that were potentially overly broad. This resulted in a lack of time to support students in the use of traditional statistical and coding techniques to analyze their quantitative and qualitative survey-based data. Despite students' frustrations in this regard, the lack of time might actually have led students to rely more heavily on collaborative data analysis, thereby foregrounding data processes as a relational rather than independent activity. The fact that the class was composed of all female-identifying students may not have been a coincidence. Some students noted the potential connection between the demographics of the class and opportunities for shared vulnerability and for raising concerns about the lack of available data for women and people from racially minoritized backgrounds.

Attuning to Noise in a World of Signal

In this section, I discuss this dissertation study's contributions to and resonances with the existing literature on data literacies/science education. One of the main contributions of this study is to build on existing efforts to teach data literacies/science in critical, creative, and innovative ways (e.g., Fotopoulou, 2021; Stornaiuolo, 2020; Van Wart et al., 2020). These efforts emphasize that data are not neutral, given, or inherently objective. Data possess a formatting power that, in part, stems from data's connection to mathematics, statistics, and computer science (Rubel et al., 2021). Rather than merely offering opportunities to critique data,

however, these efforts provide alternatives for students to "understand data as contextualized resources available for multiple purposes" (Stornaiuolo, 2020, p. 94). Such alternatives promote an expansive understanding of what counts as data, how data appears throughout our lives, and why individuals and communities might want to study data and data processes in the first place. Despite the rich body of literature emerging around the topic of critical and creative approaches to data literacies/science education, this topic remains in its early stages of development given the relatively recent, though rapid, surge in interest around data throughout all parts of society. Thus, raising new questions about and seeking new possibilities for teaching and learning about data continues to be an important, impactful, and urgent area of study because of the many ways that individuals and communities can approach this work.

In the case of this dissertation study, I interpret my unique approach to this work as one that attunes to and elevates the concept of noise as a potentially generative concept for data literacies/science education. Rhetoric around the importance of data literacies/science education too often suggests the need for and possibility of wielding data and data-driven tools to gain control or mastery over aspects of our personal lives and over issues of broader social concern (see e.g., Bargagliotti et al., 2020). Echoing critiques by D'Ignazio and Klein (2020), among others, there is a widespread belief that data can grant its wielders a birds-eye view of the world and therefore access to a supposedly objective layer of reality that is "ontologically superior to the one we actually inhabit" (McQuillan, 2018, p. 8). In Chapter 2, in particular, I characterize this script about data in terms of elevating signal over noise. The purpose of this dissertation study is not so much to flip the script and promote noise over signal but rather to enfold the two concepts together, attune to their entanglement, and consider the various ways that the intra-active dynamic between signal and noise serves as a productive metaphor for making sense of

critical and creative approaches to data literacies/science education. The following subsections unpack the contributions of this dissertation study in greater detail.

Mediation Versus Agency

Some of the literature on critical and creative approaches to data science education ground their work in Freire's literacies method of reading and writing the world (e.g., Fotopoulou, 2021) or in D'Ignazio and Klein's (2020) intersectional feminist approach to data science (e.g., Kahn et al., 2022; Lee et al., 2022). This dissertation builds on this literature by grounding work in data literacies/science education in relational ontologies and a diffractive understanding of teaching and learning within data-intensive environments. As such, each of the three chapters of this dissertation moves alongside the work of Dixon-Román (2017), McQuillan (2017), Sanches et al., (2022), and others, who think diffractively about data and data science, though not necessarily within educational contexts. There are several implications and contributions to the field that emerge as a result of taking a diffractive approach to data literacies/science education. First, while some of the literature encourages educators to create opportunities for students to encounter the "material resistances" of data (Hardy et al., 2020), Chapters 2 and 3 of this dissertation suggest that educators and instructional designers can attend to the material world as more than a mere mediating agent. The agentic capacities of a data-intensive learning environment include more than the people and discourses that occupy the space but also include the physical objects, infrastructures, technologies, and other bodies that make up processes of data generation, analysis, and communication. Matter, moreover, is not separate from discourse but are instead entangled as discursive-materials (Barad, 2007), which Chapter 2 describes in terms of noisy liminal agents and which Chapter 3 examines within its analysis of Saba and Bobble's activities with respect to their data postcard and data

physicalization activities. The discussion and findings of Chapters 2 and 3, taken together, underscore that while there is a need to more closely attune to the agentic capacities of the more-than-human world within data-intensive learning environments, matter is not always cooperative or in harmony with human goals. The concept of noise becomes a useful device for making sense of this idea. The agentic capacities of the more-than-human world may offer a degree of "material resistance," but as Serres (1982) emphasizes in his theory of the parasite, such noise is also necessary for data to generate something new beyond what is already known or what already exists.

Students appeared to respond to the materiality of the interdisciplinary data storytelling course in varied ways, a finding that resonates with but also differs from the discussion of student learning within Stornaiuolo (2020) among others. In Chapter 2, for instance, the physical characteristics of the Legos, cookies, and yarn—alongside the hands, gestures, and words of Saba and Bobble—intra-acted to reproduce, at times, conventional ideas about data as a pointer to an independently existing reality. At other times, however, materials and students came together to suggest unconventional, improvisational, and even non-representational possibilities for teaching and learning about data. Through their playfulness and novelty, such moments disrupt traditional desires to use data for purposes of elevating signal over noise and point to the possibility of embracing "choreosonic joyful noise" within data-driven processes. In this sense, Chapters 2 and 3 resonate with one another. Despite this dissertation's call for educators to seek out new possibilities for data literacies/science education by attuning to the agentic capacities of the more-than-human world, the discussion and findings of Chapters 2 and 3 emphasize that during any attempt to operationalize such call within a classroom setting, a mixture of traditional, non-traditional, expected, and unexpected orientations to data may emerge on account of the

intra-active entanglements of people, bodies, discourses, and materials. At the very least, students may experience a broadened perspective on the nature of data, as I discuss in Chapter 4. Once again, the concept of noise emerges as a productive way of making sense of this dissertation study because noise can refer not only to the messiness of data processes but also to the messiness of teaching and learning about them. In turn, theorizing messiness and ambiguity as themes within data literacies/science education becomes one of this dissertation study's contributions to the existing literature on critical and creative approaches to teaching and learning about data.

Ethics and Response-Ability

Another implication of this dissertation's diffractive approach concerns statements within the literature that educators should discuss the ethical dimensions of data with students. Such literature encourages educators to teach students that data can be biased and can be used for unethical purposes. Although this is an important first step toward moving beyond placing an exclusive focus on technical proficiency within data literacies/science education, my dissertation study joins the work of scholars such as D'Ignazio and Klein (2020) in order to pursue a deeper understanding of ethics and its relation to issues of power and the body. D'Ignazio and Klein (2020) argue that attending to data ethics is insufficient without also attending to data justice, which they characterize in terms of attending to power, striving for individual and collective liberation, and making visible the underpaid and undervalued labor that comprise the data pipeline, which is often organized along lines of race, gender, class, ability, and other social identity markers.

Chapter 2 connects these ideas to how "bodies marked as 'other'" (Thompson, 2017, p. 27) are often characterized in terms of noise. Laborers within the data pipeline, therefore,

become a form of "noise," whose influence traditional data practices seek to minimize, regulate, or control. I use Crawley's (2017) notion of choreosonic joyful noise to flip this script and show that it is in fact traditional data practices that act as "noise" and that intra-rupt other ways of knowing and being in the world. These traditional data practices promote scripts and valences of teleological progress and risk aversion that displace alternative approaches to working with data that are not only cognitive but also affective and rooted in the life experiences of people, particularly those from traditionally marginalized backgrounds. Chapter 3, then, shows what happened when I sought to bring these kinds of ideas into the classroom. Rather than echoing these ideas in exactly the same way that I understand them or resisting these ideas completely, students took up and enacted these ideas in varied ways, just as they had with respect to the materiality of data. Students were largely amenable to and eager to have conversations about the connection between data processes and issues of power, ethics, and the body. Nonetheless, just as had occurred in Stornaiuolo (2020) and to some extent Philip et al. (2013) and Philip et al. (2016), where attempts to carry out innovative lessons on data literacies/science fell short of expectations, Chapter 3 suggests moments where students' ideas about data did not necessarily align with my own emerging understandings. In contrast with some of the sentiment suggested in the literature, however, I do not interpret these moments as a limitation but rather as part of a response-able orientation to pedagogy (Bozalek & Zembylas, 2017) that centers attentiveness, curiosity, and response rather than the imposition of pre-settled ideas. Chapter 4, in some ways, reflects this commitment to response-able pedagogy. Despite the fact that student learning about data did not exactly align with my own emergent understanding about data, the methodological approach of participatory data analysis and co-authorship found in Chapter 4 represents my desire to think with rather than think in spite of students' varied perspectives about data. In this

sense, Chapter 4 stands out within the literature on data literacies/science education through the explicit involvement of the students about whose learning the chapter discusses.

Ambiguity and Improvisation

The literature on data literacies/science education emphasizes the importance of supporting students in engaging in accurate and thoughtful data analysis and communication (e.g., Bargagliotti et al, 2020; NASM, 2018). Each of this dissertation's chapters echo these sentiments in different ways. Chapter 2 acknowledges that despite its calls to attune to the dynamic intra-action of signal and noise within data processes, there are times and places for students to engage in traditional forms of data analysis that purport to yield evidence-based findings and that assign a degree of truthiness to one's data-driven conclusions about the world. Similarly, Chapters 3 and 4 consider student concerns around the issue of data bias and the need to avoid inaccuracies and misuses within processes of data analysis and communication. Rather than solely echoing the literature in this regard, however, each of the chapters also seek to add nuance to the idea of supporting students' sense-making with data. Konold and colleagues (2015), for instance, present four ways that students reason with data. They can interpret data as a pointer to a specific event, as a case value, as an indicator of frequency of a phenomenon, or as capable of communicating emergent properties of a phenomenon such as shape and center. Taken together, the three chapters of this dissertation suggest a fifth way of reasoning with data: data as a source of ambiguity. Data is a speculative endeavor that raises new questions just as much as they lead to inferences and interpretations. Chapter 2 theorizes this ambiguity in terms of the intra-play of signal and noise. Chapters 3 and 4 suggest moments when the students and I used ambiguity as a resource within the course on data storytelling. This notion of ambiguity is more than acknowledging that people can interpret the same set of data differently or that many

interpretations of data come with a confidence interval, whether explicitly stated or not. Rather, the literature discussed in Chapter 2 and the student statements and activities in Chapters 3 and 4 specify that ambiguity and contingency within processes of data analysis and communication arise on account of the complex entanglements among bodies and discourse-materials within data-intensive learning environments. Moreover, the term ambiguity encompasses unresolved questions about the epistemological and ontological foundations of data literacies/science and the connections between data and other ways of knowing and being. Embracing and attuning to ambiguity, broadly construed, therefore is a significant contribution of this dissertation study to the existing literature.

Closely connected to the idea of ambiguity is the idea of improvisation. Much of this dissertation builds on or extends ideas that exist within the literature on data literacies/science education. However, few, if not any, scholars in data literacies/science education discuss the potential role that improvisation can play within data-intensive learning environments. Just as in the case of the term ambiguity, this dissertation study suggests a broad conceptualization of improvisation that includes not only the improvisational acts of people but also the improvisational agentic capacities of the more-than-human world. Chapter 3 briefly discusses a connection between the potential for improvisation within data processes and the improvisational play of musical jamming. Chapter 4 suggests moments of improvisation when students came together to create their data stories, either by working with a yarn-based data physicalization or engaging in collaborative data analysis and communication through the creation of a podcast. Both chapters contribute to the literature on data literacies/science education by suggesting the possibility of breaking free from instructional design as a process of predetermined deliberate action, or even as a cycle of iteration and revision. Instead, teaching and learning about data

should be a matter of attunement and response. Although the literature does not explicitly mention the term improvisation, this idea of attunement and response resonates with McQuillan's (2018) call for a new materialist orientation to data science and to Dixon-Román's (2017) suggestion that the sociotechnical assemblages of data are more than the product of intentional activity but rather bear the markings of sociopolitical relations that "contaminate the haunted products of social inquiry" (p. 46). By elevating ambiguity and improvisation as legitimate forms of data reasoning, educators within data-intensive learning environments can help students attune to these "haunted products of social inquiry."

Storytelling

The themes of ambiguity and improvisation emerged alongside the idea of storytelling throughout this dissertation study. Within the literature on data literacies/science education, scholars often discuss the role of storytelling in terms of supporting students in crafting data-driven stories. Data-driven storytelling includes creating a narrative based on the results of data analysis or encoding the results of data analysis into a data representation. Chapters 3 and 4 build on this aspect of the literature in several ways. First, the chapters discuss the orientations to data that emerged throughout the data storytelling course, which had an explicit focus on supporting students in crafting data-driven stories from data that they generated with respect to themselves—as in the case of the data postcards—and that they generated with respect to a survey that they disseminated to other students. Chapter 3 highlights the role that materials play within processes of data-driven storytelling and discusses the dominant and non-dominant epistemological and ontological ideas about data that emerged as a result of telling stories through data. In Chapter 4, a key finding included the fact that students' conceptions of data broadened within the context of the data storytelling activities with which they engaged

throughout the course. Further, while much of the literature on data storytelling does not explicitly define or suggest any particular conceptualization of what constitutes a "story," Chapters 3 and 4 refer to various ideas about the meaning of a story. In Chapter 3, for instance, I refer to stories as a structured form of communication and expression and, at the risk of conflating the related but distinct idea of stories and narratives, discussed my use of Freytag's Pyramid within the data storytelling course. Bobble offered an even more expansive definition of stories as emerging out of the fact of something's existence. Importantly, Chapter 4 highlights a conceptualization of stories and data storytelling that is largely absent within the existing literature on data literacies/science education. In particular, students took up the idea of data as a form of story itself. Data are not only a way of telling a story about the world, but because data are necessarily a part of the world, then data and data processes are themselves stories that are a part of the world's becoming. This idea of data-as-story resonates with a relational ontological worldview and, in this sense, can contribute a conceptualization of data that is less common within the field of data literacies/science education.

Interdisciplinarity

The last contribution of this dissertation concerns the interdisciplinarity of data, data literacies/science, and data literacies/science education. The literature on data literacies/science education acknowledges the interdisciplinarity of the field, often in terms of the need for interdisciplinary teams of or interdisciplinary training for educators to ensure that students learn the wide range of skills, dispositions, and conceptual knowledge connected to real-world data literacies/science practices (NASM, 2018; Weiland & Williams, 2023). Less prevalent within the literature is a discussion of data-intensive learning environments as an interdisciplinary space where students can draw on and learn from one another's various disciplinary backgrounds. The

course on data storytelling discussed in Chapters 3 and 4 provides an existence proof for a course that provides opportunities for such interdisciplinarity. Chapter 4 suggests that the cross pollination of ideas, as informed by the unique personal and disciplinary backgrounds and career goals of the course participants, became a key part of students' engagements with collaborative data generation, analysis, and communication. Further, the chapter suggests the need for further refinement of the course or similar courses in order to provide even more sustained collaboration across diverse groups of students. The chapter also indicates that while the creation of interdisciplinary goals for the course may have led to a potential lack of focus or confusion around the intended takeaways of the course, interdisciplinarity may have created conditions for more critical and creative encounters with data and data processes. One important finding from Chapter 3 is that when students experience critical and creative encounters with data, a potential role for educators is to hold conversations with students about the normativity of these encounters and how they either conform or do not conform to the standards and expectations of various disciplinary communities.

Personal and Broader Impacts

I am completing this dissertation during a time when I have the privilege of stepping into a professional role where I can continue to consider the implications of this work that I have begun in graduate school. In particular, this dissertation study has forthcoming implications for my work in two settings: first, in my role as a member of a founding committee for a new U.S.-based data science education conference and second, in my role as an emerging faculty in mathematics education.

A New Data Science Education Conference

My involvement with the new data science education conference has the most direct connections to this dissertation study. Each of the founding committee members have been asked to lead a session at the forthcoming conference in February 2025. One goal of the conference is to host innovative sessions that go beyond paper and poster sessions, symposia, and working groups. Currently, plans are underway to create a broad range of session types such as interactive demos and lessons and collaborative design sessions. That said, one potential impact of my dissertation study is to share work from this study through an innovative session at the data science education conference, along with other conferences such as the AMTE or NCTM Annual Meeting. The purpose would be to share this work with data literacies/science education researchers, educators, policymakers, advocates, and other stakeholders, along with teachers and teacher educators interested in teaching and learning about data. My hope is that by next year, I will be able to develop and share one or more interactive lessons or activities that incorporate and build on the findings of this dissertation, ideally in a way that enables me to develop some of the arts-based elements of my work that I was not yet able to share in this study.

Mathematics Teacher Education and Informal Learning Environments

My role as a future mathematics teacher educator raises several questions about the intended audience and potential impact of this dissertation study. In this study's introduction, I discuss the fact that I entered graduate school with a background in secondary mathematics teaching. Although my initial research interests in graduate school was situated directly in the field of mathematics education, the data storytelling course in which this dissertation study is grounded was not a course on mathematics or mathematics education. Nonetheless, I believe that my former experience as a mathematics teacher had an inevitable impact on the course design,

on my pedagogical approaches within the course, and on several choices that I made while writing this dissertation. My involvement in a writing group that includes many mathematics education scholars, for instance, led me to read the literature on data science education that is situated within mathematics and statistics education contexts (e.g., Bargagliotti et al., 2020; Lamar & Boaler, 2021; Rubel et al., 2021; Weiland & Williams, 2023). This body literature, in turn, became important aspects of my theoretical framing, findings, and discussion throughout various parts of this dissertation. Other ideas from mathematics education, such as Martin's (2019) critique of mathematics for all and de Freitas and Sinclair's (2013) idea of inclusive materialism made an explicit appearance in one of this dissertation's chapters and influenced my thinking throughout the study. It is also important to point out that I went on the job market and received an offer for a position in mathematics teacher education during the process of writing this dissertation study, which shaped the direction of some of this study's chapters. These influences have shaped the dissertation in ways that make it appropriate not only for people who identify as data literacies/science educators but also preK-20 mathematics pre-service and in-service teachers, particularly those with an interest in incorporating data, probability, and statistics into their classrooms. The relevance of this dissertation study for mathematics educators will only grow stronger if and when states begin to adopt policy changes that formalize the expansion of data literacies/science offerings within elementary and secondary mathematics curricula. In New York, where I will be heading to next year, there have been no prominent, public discussions of incorporating data literacies/science into preK-12 mathematics curricula in the same way that has been discussed in states such as California. Nonetheless, I intend to incorporate my developing expertise in data literacies/science education into my position as a mathematics teacher educator, either through coursework explicitly designed to teach data,

probability, and statistics to pre-service mathematics teachers or by introducing examples into content courses that ask my students to think about topics in data literacies, data science, probability, and statistics.

The dissertation study also has relevance for informal learning environments. I have had the opportunity to be involved in the preparation of a research grant connected to work with youth and youth leaders in a local subsidized housing community. Part of the work occurring in this space involves understanding the data practices of youth, including how they make sense of, communicate, and use data in ways that are not necessarily prescribed by adult researchers. This dissertation study's discussion of the use and agentic contributions of the material world in connection with data practices, along with the intra-actions of signal and noise that emerge when young people are given opportunities to engage with data in critical and creative ways, are potentially relevant to the work of educators and youth within such informal settings. The dissertation's impact, therefore, can reach beyond the walls of the classroom, especially given the ubiquity of data-driven technologies that can lead to no shortage of conversations and activities connected to critical and creative data use.

Theory and Teaching

This dissertation study draws heavily on theory, whether through the development of theory in Chapter 2, the use of relational ontologies in Chapter 3, or the use of participatory methodologies such as the voice-centered relational method in Chapter 4. The dissertation, along with my work throughout graduate school, leans on Jackson and Mazzei's (2011) notion of thinking with theory as an orientation to teaching and research. In writing this dissertation, I am now considering what it will mean to "bring" this study's theoretical work to my future pre-service mathematics teacher students. I place the term "bring" in quotes because the theories that

I have begun to develop here are not settled and, therefore, the ideas that I will "bring" to them cannot be likened to a pre-developed object. I am reminded of a statement that my former Shakespeare professor Michael Collins repeatedly said while I was enrolled in his class as an undergraduate student: "There is no there there." By this, he meant that the written text of any Shakespearean play—or any play, generally—does not constitute the play. The play only becomes a play once it is interpreted and enacted by a particular group of stage performers, to a particular audience, in a particular place, at a particular time, and enmeshed in a particular ocean of materials. In other words, there is no play until it becomes one. In a similar manner, this dissertation study does not constitute a standalone set of ideas with immediate implications for or impact on fields such as teacher education. Instead, when I "bring" ideas about relational ontologies of data or data as a form of story to my students, these ideas will only truly emerge as they are shaped, interpreted, constructed, taken up, challenged, and ultimately formed into being through the mutually constitutive activities of all agentic participants in the learning environment. Indeed, the idea that "there is no there there" gestures toward—and is likely connected to my initial interest in—a response-able orientation to pedagogy grounded in rejecting the idea of pre-configured people, objects, and ideas.

I believe that theory is an important component for teacher education and the everyday work of teaching, and I interpret this dissertation study as a tacit rejection of the binaries too often expressed between theory and practice and between teachers and researchers. I have learned in graduate school that how we live our lives is always shaped by and shaping our theories about the world. Scholarship in feminist new materialisms, for instance, has taught me that theorizing is not merely the production of abstract ideas. Instead, theories impose a performative effect on our reality and is part of "the world's worlding itself" (Barad, 2011, p.

133). I often return to the following quote from Haraway (2016): “It matters what thoughts think thoughts. It matters what knowledges know knowledges. It matters what relations relate relations. It matters what worlds world worlds. It matters what stories tell stories” (p. 35). Her words not only apply when we conduct research but also when we teach. What stories are we telling to tell stories about our students? About ourselves as educators? About the content that we teach? About the broader sociopolitical context within which we teach? I intentionally chose to conduct a dissertation study in which I was simultaneously the teacher and the researcher in the same space in order to attune to the entanglement between both forms of knowledge production. From my engagement in this dissertation study, I have learned that when teachers strive to enact otherwise possibilities for teaching and learning, they are necessarily positioned as researchers and theory-builders. This is because both reform-oriented teaching and research are matters of either seeking the new or re-surfacing that which already exists. Indeed, in this study, along with many other research projects with which I have been involved during graduate school, I have found that educational research often requires the researchers to carry out thoughtful pedagogical practices, whether it is during moments of working with other educators, working with students, interviewing participants, or helping design or carry out classroom lessons or activities within informal learning spaces. Further, when teachers seek to enact reform-oriented teaching practices, they can consider the relevant literature, develop an explicit or tacit set of principles to guide their work, consider their positionality and the broader context of the learning space, attune and respond to what emerges in the learning space, and share their work with others. In these ways, teaching and learning are intertwined, and this entanglement was no less present as I planned, taught, considered, and continued to work with the students in the data storytelling course. It is therefore my hope that how I sought to think with theory during my time as a

teacher-researcher engaging in teaching-research can resonate with and even inspire educators to deeply engage with theory in their own contexts as part of the broader work of equitable and justice-oriented instruction.

What Next?

There are many ideas, questions, and tensions that arose throughout the course of teaching the data storytelling course that I was not able to address in this dissertation study and that will constitute research that I plan to consider, formally or informally, moving forward in my career. In this section, I briefly share these issues, which in the spirit of this dissertation raises more questions than provides answers. Upon entering the seminar room on the first day of the data storytelling course, I was immediately struck by the fact that I was a Filipino American man in a space that was predominantly, though not exclusively, made up of white women. The course demographics reflected the racial and socioeconomic demographic characteristics of the primarily white institution in which the course was taught. Given that my interests in data literacies/science education have concerned the disruption of dominant epistemologies and ontologies of data that are rooted in colonialism and that cause harm across lines of race, gender, class, ability, and other social identity markers, a question that arose on that first day of class was how the context of my study would impact the goals and interests that I had in mind for re-imagining data literacies/science education. Surely, it would be naive to say that justice-oriented work cannot be enacted in a primarily white space. But it would be equally naive to say that such context is irrelevant. In fact, several moments arose in class that I did not discuss in this dissertation that caused me to confront how my positionality and beliefs about teaching and learning about data intra-acted with my students' responses to and engagements with some of the activities that I designed. These activities included, for instance, conversations around the ethical

uses of data given the historic origins of modern statistical thought and its close connections to the eugenics movement. It stood out to me how some of the student responses paralleled and, in some cases made explicit reference to, debates surrounding the separation of art from the actions and behaviors of the artist. At the time, I had asked myself and continue to ask myself today: were these adequate responses? How should I respond or have responded to what my students said? Moreover, what does it mean that I was a man (again, a Filipino American man) in an otherwise exclusively female space, talking about a typically male-dominated subject? I believe that many of the instructional choices that I made throughout the course—including the use of Data Feminism (D'Ignazio & Klein, 2020) as a guiding text; the references to artists such as Mona Chalabi; the conversations around data, power, ethics, and the body; and the elevation of other ways of knowing and being beyond the boundaries of disembodied reason—speak to my positionality in relation to those of my students. Many of these instructional choices, in turn, can be found within Chapters 3 and 4 of this dissertation study. Nonetheless, I did not go into adequate depth about these issues in this study for fear that this study's scope would spiral out of control. For this reason, my hope is to continue to consider the impact that my identity played in the teaching and research space, as well as the impact that my identity will play with my students next year.

In this conclusion, I state that this study provides an existence proof for a course unconstrained by the limitations of disciplinarity. It is also true, however, that the course was made possible because it counted as a pass/fail Honors College course. How did this listing as an Honors College course shape the outcomes of the course and of this dissertation study? How might the ideas and findings of this study apply to course contexts that do not share similar freedoms of course design and instruction? These are important questions that are merely

touched on in a few places throughout the dissertation and therefore merit further consideration. One issue that this dissertation raises is how its findings relate to traditional statistics classrooms, where the disciplinary concerns of the field constrain the kinds of activities and learning goals that can take place. Despite the informal approach to data analysis that occurred throughout the data storytelling course, it could be argued that the students nevertheless engaged in deep and rigorous statistical thinking. For instance, how the students considered their initial beliefs about their chosen social issue topics and then revised those beliefs based on the survey data was arguably a form of Bayesian reasoning that entailed updating their prior beliefs using available evidence. Moreover, the idea of abductive reasoning could be likened to an informal approach to maximum likelihood estimation. More broadly, one can consider not only how educators can adapt this study's findings to more constrained classrooms, but also how educators, researchers, policymakers, and other stakeholders can work together to rework educational systems in order to enable greater interdisciplinarity. How do opportunities for interdisciplinarity cut across lines of race, class, space, and place? Is it the case that innovative, interdisciplinary learning most often occurs within certain privileged spaces? Are there spaces led by or composed of communities of color and people from other historically marginalized backgrounds where interdisciplinary learning occurs but in unacknowledged ways? These questions require serious consideration and are questions that I would be interested in pursuing in my capacity as a teacher educator.

Before coming to MSU, I had not known that arts-based research existed. If I had attended another university for my doctorate, my ignorance may have continued. However, at MSU, I learned about the valuable intersection of research and the arts through faculty and graduate student peers such as Lynn Fendler, Higinio Dominguez, Sandro Barros, Vivek

Vellanki, Liv Furman, Hannah Grisham, Renée Wilmot, Reyila Hadeer, Sofia Abreu, Darshana Devarajan, Kasun Gajasinghe, Joel Berends, Brady Tyburski, Anthony Dickson, and many others. In turn, this led me to want to explore my own visual and sonic artistic practice—particularly in the realm of data-driven art and art about data—and how it might inform some of the issues, questions, and tensions that I raise in this section. By the time of this writing, however, I felt underdeveloped in my artistic identity such that I did not have confidence that I would be able to produce and articulate the significance of my art with the sufficient quality and care necessary to meaningfully contribute to the field of data literacies/science education, to honor past traditions of arts-based research, and to meet my personal expectations as an emerging artist. Therefore, although I had prepared some arts-based pieces for this dissertation study, I have chosen to omit them from the study entirely. It is my hope that my future work in data literacies/science education, as well as my work in mathematics teacher education, can involve the incorporation of artistic practices, research, and pedagogy in ways that align with my commitment toward a relational ontological orientation to knowing and being.

This dissertation study leans on noise as a way to reconsider data literacies/science education. The fact that noise is a topic lying at the intersection of philosophy and sound studies, among other fields such as the digital humanities and education, is no accident. Throughout my time in graduate school, I became interested in the relevance of philosophy for approaching the work of mathematics teaching and mathematics education research. In the later half of my graduate studies, influenced by feminist new materialist scholarship that eschewed the overemphasis on sight and vision as central metaphors for knowledge production, as well as scholars such as Shannon and Truman (2020) and Gershon (2017) who work within the area of sound studies in education, I became increasingly interested in the use of sonic metaphors and

actual sonic data to make sense of educational systems, particularly as they pertain to issues of power, privilege, and oppression within formal and informal learning environments. In turn, this interest in sound studies in education rekindled my interest for musical performance, which I had developed as a child but set aside as an adult. The choice to center my dissertation around the theme of noise, therefore, both inspired and was inspired by the coming together of these forces. However, because my development within the sonic arts exists merely in its emergent stages, this dissertation signifies a beginning step in my journey of exploring what lies at the intersection of sound and education. Moving forward, this is one area of growth that I am enthusiastic about pursuing.

Related to the sonic arts is the notion of improvisation. The idea of improvisation appeared briefly in this dissertation study and is a topic that I am eager to consider further because of the concept's close relation to ideas of sound and noise. Near the time of writing this conclusion, I attended a roundtable session at the Annual Meeting of the American Education Research Association on the topic of improvisation (Marin et al., 2024; Rios, 2024; Sanchez et al., 2024; Sherry-Wagner & Bang, 2024; Vossoughi et al., 2024) The session was popular, with three concentric circles of non-presenting participants in attendance. During the roundtable, the presenters conceptualized improvisation as an emancipatory practice. Drawing on jazz, mariachi, multi-vocal storytelling, and other improvisational practices, they sought to disrupt the binary of structure versus improvisation and instead think of the importance of the joy of spontaneity within educational settings. Inspired by this roundtable session, my follow up work in the areas of the sonic arts and improvisation would be to further theorize, study, and surface teacher and student practices of improvisation, ambiguity, and intra-ruption. Because these are not typically emphasized within data literacies/science education, and because they in some ways challenge

norms of control and prediction that are connected with the colonial roots of modern statistical thought, I believe they are particularly ripe concepts for re-imagining teaching and learning about data within classrooms and other learning environments.

Toward a Speculative Orientation to Teaching and Learning About Data

Cajete (2000) writes, "The modern obsession of being in control and the dream of eliminating uncertainty through control of nature, which is the underlying philosophical premise of Western science, must give way to the reality of moving creatively with the flow of events, which is the true reality of the universe" (p. 16). I resonate with his words in the way that I have approached this dissertation. My interest in and use of noise throughout this study was motivated by my desire to resist the logics of control that too often pervade how we think, teach, and learn about data. This dissertation was an attempt to gesture towards an alternative orientation to data, all while wrestling with what emerged in practice during the time I spent designing and teaching the data storytelling course and collaborating with my co-instructor and students afterward. Much of my work has relied on and would not have been possible without the guidance of the people in my life, as well as the wisdom of the scholars who write about/with relational ontologies and the innovative ideas of my students. I am deeply indebted to them, and this dissertation is a reminder that all scholarly work is the product of collective striving and action.

I consider my attempts to strive toward an alternative orientation to data as a counter-speculative endeavor. Data processes are already speculative insofar as they seek to quantify, measure, and control past, present, and future uncertainty. Following Bahng (2018), I have sought to ask "if another mode of speculation is possible, one that is not immediately captured by the anxious gatherings of risk" (p. 4). Consistent with a speculative orientation to knowledge production, I embrace that this dissertation's attempts at seeking out such other modes of

speculation were neither tidy nor conclusive. Instead, I argue that the goal of this study, as is the goal of projects within the field of speculative design, was "to create spaces for discussion and debate about alternative ways of being, and to inspire and encourage people's imaginations to flow freely" (Dunne & Raby, 2013, p. 2). To this end, I conclude this dissertation by briefly engaging in a moment of speculation. Inspired by the imaginative possibilities generated by the students in the data storytelling course, I offer a brief glimpse into yet another possibility for what practices of data visualization can look like beyond the confines of scatterplots, bar graphs, and histograms traditionally taught in school settings (see also Lim et al., 2022 for a discussion of a similar topic). I ask the following: What if our forests could "wear" the clothes and products that the beauty industry associates with what is "natural" and "desirable"? What might they teach us? Figure 13 ("Makeup Forest") is an experiment in data visualization based on data collected from "The Naked Truth," a 2020 study by Ofunne Amaka and Amber Thomas to examine how 6,816 complex products and the names associated with them reveal bias in the beauty industry. I used the data to produce the visualization using the ggplot2 library in R, a programming language used for statistical computing and graphics. The characteristics of the trees are bound to features of a particular company's product offerings—for instance, each leaf is randomly generated from a selection of makeup shades sold by that company. The figure invites readers to consider one possible future for teaching and learning about data. In this imagined future, data literacies/science education is a space where data, art, nature, and broader sociopolitical issues meet. The sharp distinctions between what are and what are not data become blurred. The purpose of using data is not to arrive at settled conclusions but rather to raise questions that are simultaneously grounded and fantastical in nature. Data-intensive learning environments become spaces of speculation where students can use data to seek out otherwise possibilities (Green,

2020) for understanding themselves and their world. Data takes on a critical but also agentic role, collaborating with students to address social issues such as those that arise at the intersections of race, gender, aesthetics, and capitalism. The possible future that Figure 13 considers is one among many and raises the following question: What other forms of data literacies/science education are possible?

Figure 13. Makeup Forest

I end with the words of adrienne maree brown (2017) in her book *Emergent Strategy: Shaping Change, Changing Worlds*. She writes, "Some of us are surviving, following, flocking—but some of us are trying to imagine where we are going as we fly. That is radical imagination" (p. 21).

REFERENCES

Amaka, O., & Thomas, A. (2021). The naked truth. *The Pudding*. Accessed at: https://pudding.cool/2021/03/foundation-names/

Bahng, A. (2018). *Migrant futures: Decolonizing speculation in financial times*. Duke University Press.

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.

Barad, K. (2011). Nature's queer performativity. *Qui Parle, 19*(2), 121-158.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D.A. (2020). *PreK-12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education*. American Statistical Association and National Council of Teachers of Mathematics.

Bozalek, V., & Zembylas, M. (2017). Diffraction or reflection? Sketching the contours of two methodologies in educational research. *International Journal of Qualitative Studies in Education, 30*(2), 111-127.

brown, a. m. (2017). *Emergent strategy: Shaping change, changing worlds*. AK Press.

Cajete, G. (2000). *Native science: Natural laws of interdependence.* Clear Light Publishers.

Carroll, S. (2024). TEDxUArizona: Indigenous Peoples Breathing Data Back. *TEDx Talks*. Accessed at: https://www.youtube.com/watch?v=jPS_3mZXWXw

Chen, G. A. (2023). Teacherly response-ability: ethical relationality as protest against mathematical violence. *Educational Studies in Mathematics*, 114, 277-296.

Crawley, A. T. (2017). *Blackpentecostal breath: The aesthetics of possibility*. Fordham University Press.

de Freitas, E., & Sinclair, N. (2013). New materialist ontologies in mathematics education: The body in/of mathematics. *Education Studies in Mathematics, 83*, 453-470.

de los Rios, C. V. (2024). "Convivencia not competencia": Learning from an elementary-aged girl mariachi ensemble. Roundtable session at the *Annual Convention of the American Educational Research Association*. Philadelphia, PA.

D'Ignazio, C., & Klein, L. (2020). *Data feminism*. The MIT Press.

Dixon-Román, E. (2017). Toward a hauntology on data: On the sociopolitical forces of data assemblages. *Research in Education, 98*(1), 44-58.

Ellenwood, C., & Foxworth, R. (2024). Data back: Indigenous material culture, philanthropy, and Indigenous data sovereignty. *The Public Historian, 46*(2), 123-143.

Fotopoulou, A. (2021). Conceptualising critical data literacies for civil society organisations: Agency, care, and social responsibility. *Information, Communication & Society, 24*(11), 1640-1657.

Gershon, W. S. (2017). *Sound curriculum: Sonic studies in educational theory, method, & practice*. Routledge.

Green, K. L. (2020). Radical imagination and "otherwise possibilities" in qualitative research. *International Journal of Qualitative Studies in Education, 33*(1), 115-127.

Haraway, D. (2016). *Staying with the trouble: Making kin in the Chthulucene*. Duke University Press.

Hardy, L., Dixon, C., & Hsi, S. (2020). From data collectors to data producers: Shifting students' relationship to data. *Journal of the Learning Sciences, 29*(1), 104-126.

Ingold, T. (2011). *Being alive: Essays on movement, knowledge and description*. Routledge.

Jackson, A. Y., & Mazzei, L. (2011). *Thinking with theory in qualitative research: Viewing data across multiple perspectives*. Routledge.

Kahn, J. B., Peralta, L. M., Rubel, L H., Lim, V. Y., Jiang, S., & Herbel-Eisenmann, B. (2022). Notice, wonder, feel, act, and reimagine as a path toward social justice in data science education. *Educational Technology & Society, 4*(25), 80-92.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics, 88*(3), 305-325.

LaMar, T., & Boaler, B. (2021). The importance and emergence of K-12 data science. *Phi Delta Kappan, 103*(1), 49-53.

Lee, V. R., Pimentel, D. R., Bhargava, R., & D'Ignazio, C. (2022). Taking data feminism to school: A synthesis and review of pre-collegiate data science education projects. *British Journal of Educational Technology*, *53*(5), 1096-1113.

Lim., V. Y., Peralta, L. M., Rubel, L. H., Jiang, S., Kahn, J. B., & Herbel-Eisenmann, B. (2022). Data visualizations for mathematics education in times of crises: Engaging the body and emotions through interactivity, multimodality, and narrative. *ZDM, 55*, 109-118.

Marin, A. M., Lopez, B. Y., Lindberg, L. E., & Dave, S. (2024). Musical improvisation and the co-creation of everyday pedagogical spaces for social change. Roundtable session at the *Annual Convention of the American Educational Research Association*. Philadelphia, PA.

Martin, D. B. (2019). Equity, inclusion, and antiblackness in mathematics education. *Race, Ethnicity, and Education, 22*(4), 459-478.

McQuillan, D. (2018). Data science as machinic neoplatonism. *Philosophy & Technology, 31*(2), 253-272.

National Academies of Sciences, Engineering, and Medicine (NASM) (2018). *Data science for undergraduates: Opportunities and options*. The National Academies Press.

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction, 34*(4), 361-388.

Philip, T. M., Way, Winmar, Garcia A. D., Schuler-Brown, S., & Navarro, O. (2013). When educators attempt to make "community" a part of classroom learning: The dangers of (mis)appropriating students' communities into schools. *Teaching and Teacher Education, 34*, 174-183.

Rubel, L., Nicol, C., & Chronaki, A. (2021). A critical mathematics perspective on reading data visualizations: Reimagining through reformatting, reframing, and renarrating. *Educational Studies in Mathematics, 108*, 249-268.

Rubel, L. H., Peralta, L. M., Herbel-Eisenmann, B., Jiang, S., Kahn, J. B., & Lim, V. Y. (2021). Theorizing data science education: An intersectional feminist perspective on data, power, and "playing the game". In *Proceedings of the 43rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA)*. Philadelphia, PA.

Sanches, P., Howell, N., Tsaknaki, V., & Helms, K. (2022). Diffraction-in-action: Designerly explorations of agential realism through lived data. *CHI Conference on Human Factors in Computing Systems (April 29-May 5, 2022)*. ACM. New York, NY.

Sanchez, I. M. G., Giuliano, A., Gzehoviak, J., Meza, J., & Westphain, K. (2024). Immigrant youth's (re-)imagining of communal well-being through multivocal, improvisational storytelling. Roundtable session at the *Annual Convention of the American Educational Research Association*. Philadelphia, PA.

Serres, M. (1982). *The parasite*. The Johns Hopkins University Press.

Shannon, D. B., & Truman, S. E. (2020). Problematizing sound methods through music research-creation: Oblique curiosities. *International Journal of Qualitative Methods, 19*, 1-12.

Sherry-Wagner, J. D., & Bang, M. (2024). Designing to support "wondering with": Disciplined improvisation and ethical pedagogical mediation within field-based inquiry. Roundtable session at the *Annual Convention of the American Educational Research Association*. Philadelphia, PA.

Stornaiuolo, A. (2020). Authoring data stories in a media makerspace: Adolescents developing critical data literacies. *Journal of the Learning Sciences, 29*(1), 81-103.

Thompson, M. (2017). *Beyond unwanted sound: Noise, affect and aesthetic moralism*. Bloomsbury Publishing USA.

Van Wart, S., Lanouette, K., & Parikh, T. S. (2020). Scripts and counterscripts in community-based data science: Participatory digital mapping and the pursuit of a third space. *Journal of the Learning Sciences, 29*(1), 127-153.

Vossoughi, S., Morales, M., Franklin, M., Carroll, S., Lansana, O., James-Harris, T., & Oumarou, D. (2024). Pedagogical weaving and collective teaching in a STEAM storytelling program. Roundtable session at the *Annual Convention of the American Educational Research Association*. Philadelphia, PA.

Weiland, T., & Williams, I. (2023). Culturally relevant data in teaching statistics and data science courses. *Journal of Statistics and Data Science Education*, 1-16, https://doi.org/10.1080/26939169.2023.2249969