This Is Not a Pipe: Essays on Man as the Measurer of All Things

An Honors Thesis (HONR 499)

by

Rebecca L. Jackson

Thesis Advisor

Paul Ranieri

Signed

Ball State University
Muncie, Indiana

August 2015

Expected Date of Graduation

July 2015
Abstract

Fields of inquiry are differentiated by the methods, metrics, and conceptual "vocabulary" that define the way one approaches a problem and what counts as a solution. Yet, if different methods of measuring are what separate and distinguish various disciplines and ways of knowing, then measurement is also the uniting feature underlying all human pursuits towards understanding. In an effort to explore the nature of measurement itself, I created *This Is Not a Pipe: Essays on Man as the Measurer of All Things*, a multidisciplinary collection of essays authored by Ball State University professors of history, mathematics, theatre, literature, philosophy, linguistics, creative writing, economics, sociology, humanities, and computer science. In addition to this booklet, I include a post-project analysis which reflects on how the essays further illuminated (or contrasted with) my own pre-project hypotheses on what it means to measure.
Acknowledgements

I owe tremendous gratitude—not only for the successful completion of this project, but also for the incredible journey that has been my undergraduate education—to a long list of amazing humans. An enthusiastic thanks to:

Each participant who took time out of her or his own hectic research agenda to contribute an essay for this project—some of whom barely knew me, and several of whom are dear friends and longtime supporters of my madness.

My thesis advisor, Paul Ranieri, whose relentlessly peaceful outlook (even in the midst of my utter thought-chaos) has proved an essential friend and mentor throughout this experience, patiently encouraging me on, reigning me in, and keeping me going.

My design gurus, Rebekah Hobbs and Michelle Avelis, for their design expertise, advice, and patience.

My friends Liz and Dan, who were really the only people I ever saw during my last crazy year at Ball State, and were really all that I needed.

Jason Powell, for the many arguments and laughs; Todd McKinney, for introducing me to Creative Nonfiction; and both of them, for letting me borrow a small library’s worth of their books for a year.

Barb Stedman, for being Barb Stedman. There will only ever be one of you, and when you retire, Ball State will have to hire three people to even come close to equalling all that you do for your students.

Jim Ruebel, for teaching me that character is more important to maintain than a perfect GPA, and for sharing my conviction that, eventually, “Everything is everything else.”

John Emert, for encouraging and supporting me even through some of my more over-ambitious ideas, and for discouraging my attempts to wean myself off of a caffeine addiction. After all, “A mathematician is a machine that turns coffee into theorems.”

Rich Stankewitz, (1) for not giving me the “F” I probably deserved on the discrete systems paper he assigned to be about logical communication, and which I wrote mostly about communism (via logic notation), and (2) for every philosophical, mathematical, tangential conversation we’ve had since.

Brent Blackwell, for introducing me to the word “interdisciplinarian,” and for introducing me to myself as someone with valuable ideas worth pursuing. Without his continual support and encouragement throughout the last three years, I doubt the idea for project would have ever existed, let alone have taken flight.

My family, whose continual support, teasing, and unconditional love have made me who I am today. Particularly my parents, who have never ceased to amaze me. Their constant prayers throughout my every wandering exploration, failure, or success have been essential to my growth and intrepid ability to put one foot in front of the other.

Most of all, the One whose measures are not our measures—to whom every sparrow counts.
Table of Contents

Artist Statement ________________________________________________________________ 4

Collection of Essays - *This Is Not a Pipe: Essays on Man as the Measurer of All Things*

Post-project Analysis __________________________________________________________ 14

Appendix: Project Invitation ____________________________________________________ 47
Artist Statement

Explanation

“No, no! The adventures first, explanations take such a dreadful time.”
- Alice in Wonderland

The following booklet contains a compilation of essays written by thirteen Ball State University professors, along with my own contribution. The scholars who appear in this collection each received an invitation, asking them to ponder the idea of measurement and create an essay response to any of the following four prompts:

1. What is the role of measurement in your field?
2. What does it mean to measure with accuracy?
3. Describe a metric within your field, or perhaps the standard “ruler” you use.
4. Describe a problem with taking good measurements, or a problematic of effect of taking measurements.

Inspiration

/ˌinspəˈrāSH(ə)n/
noun

1. the process of being mentally stimulated to do or feel something, especially to do something creative.
2. the drawing in of breath; inhalation.

Creating this booklet was my attempt at salvaging the countless hours I’ve had stolen from me in the last two years, during which my attention was shanghaied into a relentless fascination with the concept of measurement.

It was cropping up everywhere—not just in my math classes—in my lyric essay class, in my class on the turbulent sixties, at Thanksgiving dinner, in linguistics, in wedding ceremonies. Questions about what it means to measure, the effects of measurement, and how to define measurement accuracy became the backdrop of every area of my education. Methods and metrics in various areas of my interest began conversing with one another about these questions, and there my trouble really began. Many late nights I fell asleep pondering how fractal geometry could help us understand how to better evaluate student progress. Other days, I woke up early with some new
pressing realization: an event in history could become a kind of “word” in discourse, operating as an approximative representation of a larger idea, and by this action of reference, changing how the event itself was later perceived. Or is this all nonsense? These half-baked epiphanies were especially problematic because of their reluctance to fit into a single category of inquiry. The natural “Who should I talk to about this,” question became a deeper one, with a clear but complex solution: “How many do I need in order to talk about this?”

My professors have always been willing to humor my questions that pertained to their particular field, but only a handful were comfortable venturing beyond that scope, to the liminal space I was wandering further into. This thesis project was the result of my compulsion to wrangle various perspectives on measurement together in an attempt to tell a broader, more comprehensive story of humankind as measurers. In addition to developing out of my own personal investigation of this topic, this project was also a scheme to assemble some of the best minds I knew into the same “room,” so to speak. The booklet becomes a discursive space where the essays can speak to each other, even if this cooperation is inadvertent. Each essay illuminates the others, surfacing new questions through the unexpected connections it makes with its neighbors.

These pieces create a collage of knowledge and ways of knowing, all centered around the role of humankind as measurers. Thus, as a whole, the collection acts as a kind of Rosetta Stone, with the “language” of each perspective on measurement adding to the understanding of the others. This allows the reader to bounce back and forth between the foreign and familiar, using her own previously acquired knowledge to piece together an understanding of unfamiliar vocabulary or calculations. It is not new or profound to state that what distinguishes the boundaries between academic areas of study are the methodologies, metrics, framings, and terminology specific to that discipline. The lucidity of this fact overshadows a broader truth: If different methods of measurement are what separate and distinguish various disciplines, and ways of knowing, then measurement is also the uniting feature underlying all human pursuits toward understanding.

There has never been an age with more information (or more access to it) than the one in which we now live. There has never been more specialized research set into motion to meet the demands brought on by a globalized age of complex problems—problems that are, across all cultural and disciplinary borders, shared. If we cannot avoid sharing problems, we cannot afford not to share solutions. My hope is that this collection draws attention to the commonalities between areas of study, blurs the lines between the goals and struggles of divergent fields, and creates a dialogue among different methods of measuring.
Process

“Building a system is the process of learning how to build the system.”
- Paul Gestwicki

The first step in this process was creating a standard invitation to send out to professors I thought might be interested in participating in the project (see appendix). Although the invitation was standardized, the body of each email was personally tailored to the receiver. In addition to sending a PDF of the invitation via email, I decided that (if at all possible) I would make appointments to meet with each one to explain the project in person. This allowed me to share my enthusiasm face-to-face, a far more persuasive and personal method—and more fun for me. Above all, I knew putting forth this extra effort kept me (and my project) from being just one more email traveling to the bottom of the inbox.

I sent invitations out first to those I knew personally, whom I had already casually approached with the idea. I needed to get a few “yes”-es under my belt, in order to build up my confidence. The majority of these were professors I had taken classes from or had known from working at the Honors College. After receiving affirmative responses to all but two (due to medical leave), I branched out to ask those whom I was acquainted with or who had been recommended to me for the project. There was also a somewhat sneaky strategy to this order-of-operations: I anticipated that those unfamiliar with me as a student would be more likely to contribute my project if they knew colleagues who were already participating.

Collecting together all the observations about measurement I’d scribbled down during the past couple years, I typed up a document containing all my pre-project hypotheses (along with examples) about measurement. It was important to me that I record them before I saw any of the contributors’ essays, so that I could later revisit the hypotheses and write a reflection on what I had learned about each one since reading the complete the booklet.

Meanwhile, I kept a schedule of reminder emails. I let each participant arrange a deadline and schedule a reminder email from me beforehand, at their convenience. For many, it took 2-3 follow-up emails to finally secure their contribution. As the essays began to roll in, I found myself overwhelmed with gratitude for how much support I felt from the authors. I could only take in one essay at a time, it was such an intense experience. Reading their thoughts on measurement, which often mirrored (and added to) my own, was amazingly rewarding and reassuring.
Finally, after weathering a few rejections, abating some skepticism, and celebrating the many individuals who had chosen to invest themselves in the project, I was ready for a risk. I crafted one last handful of emails to the professors I was least connected with, which were met by mostly positive responses. This was surprising at first, but could be attributed to my inclusion of a running list of participants in the body of the email, a list which was now quite substantial. My original goal was to secure responses from 12 participants (as I told my advisor at the start of the project: “I'd really like a dozen, but I'll settle for twelve.”). Though this was an ambitious goal, I ended up exceeding it, reaching a total of 13. The final roster includes professors from History, Mathematics, English, Economics, Sociology, Honors Humanities, Computer Science Departments.

Of course, at this point in the project, having reached my goal, I immediately set a higher one. Instead of twelve, I now wanted eighteen essays. And, I giddily decided one day, I was going to create artistic responses to each of the essays, because I had felt so strongly that I needed a way to respond to their ideas in a way that highlighted some aspect I found meaningful, and I wanted there to be more visual elements to the booklet. In addition, I saw several holes I wanted to fill. I had yet to secure a response from a physicist, a musician, an artist, or a psychologist, all of which I thought to be essential to a well-rounded study of the measurement concept. Yet, as the semester was coming to a close, final projects in other classes began crowding out the time I had available to dedicate to my thesis. This, along with the advice of my thesis advisor, prevented me from continuing on in an endless cycle of revisions and additions to the project, a tempting (and undoubtedly never-ending) chase.

Design
/dəzəˈdʒən/  noun
1. purpose, planning, or intention that exists or is thought to exist behind an action, fact, or material object.
2. the ultimate source of frustration, second-guessing, and endless revision/despair

I managed to let go of some of my newly-acquired aspirations for the project, however, I did succumb to my urge to completely overhaul the first cover design. The original cover combined the smoking pipe cross-section and title typography into a logo-like, centered image on a dark blue background. As beautiful and clean as this design was, it did not quite have the "sketched
blueprint" feel that I had originally intended. I wanted the pipe image to appear more hand-drawn, to instantly give the impression of being a sketch of measurements/dimensions for a pipe, rather than just a white outline of pipe. Thus, the final cover design appears to be made of blue graph paper, with two different perspectives on the pipe to emphasize the intended allusion toward blueprint drafting. The cover design is meant to allude to the painting "The Treachery of Images," more famously known for its French caption which translates to "This is not a pipe." The painting points out the overlooked obvious: there is a difference between a two-dimensional representation of an object and the actual object itself. Similarly, I wanted the cover to allude to the notion that measurements of an object do not equal the object itself, a distinction that is referenced in various ways throughout the booklet.

I designed the booklet's interior using InDesign, with a mostly minimalist look. The only decorative elements are the sketchy line segments frames the title text of each essay, as a way to clearly demark the beginning of each new essay. The lighter and darker lines that make up each frame suggest the act of approximation, as if the lines were different attempts to get closer and closer to forming the rectangular frame around the title text. On the original design for the back cover, the names of all of the authors were collaged, hashing together to give the impression that the names are somehow clashing and combining simultaneously. This was meant to depict how the ideas and perspectives of the authors, being placed beside one another, have opportunity to create contrast and collaboration. Yet, when I changed the cover background to blue graph paper, there was too much fighting between the lines of the background and the hashing of names, to the point where the design was looking more confusing and frustrating than clean and interesting. When I talked with my sister (who does this kind of thing professionally) about my design struggles, we came up with the idea of presenting the names along an axis, with page numbers accordingly, like a table of contents on a ruler. I think this design looks much better, but I still have a disquieted voice inside that reminds me that great design should be focused on an interpretive feel that you want the reader to experience from reading the booklet, not just on making obvious visual references to the subject of measurement. But, sometimes you have to pick your battles when it comes to design, and not everything that works in the mind will work as well on the page. Ultimately, you have to go with what works on the page.

For future projects, I would be better off starting out with two or three cover page prototypes that I develop simultaneously, eventually choosing one for the final product. I think this

---

would have allowed for more flexibility early on, avoiding some time-consuming battles between my design wish-list and the cover's aesthetic limitations.

---

**Re-visions**

"The first sentence can't be written until the final sentence is written."

- Joyce Carol Oates

My original plan of attack was to write down all of my pre-project hypotheses about measurement, read the responses, and write a reflection about how they related to or contrasted with my hypotheses. Then, from that reflective process, with additional outside research, I was going to write an essay about what it means to measure. I wasn't sure whether this essay would be a part of the booklet of not, but I suspected that it might be too long to include.

I ran into a problem with this plan early on: I had expected to have about 5 pages of hypotheses to start off with. When I finished getting down all the ideas I had been accumulating in the past couple years, the document was closer to 25 pages. Needless to say, I had underestimated how much work it would take to go through my hypotheses and compare them to what was talked about in the essays, let alone the intense time it would take to coalesce all of the resulting reflections into an essay that somehow pretended to sum up everything I had learned. Then, there was also the fact that I had done only about a third of the outside research I had hoped to do for the final essay. Once again, if it weren't for the wise advice of my thesis advisor, who encouraged me to scale down the project to something more manageable, I probably would have been too stubborn to admit I needed to adjust my expectations for the completed project.

About a year prior, I had heard about a conference called "Measuring Culture," hosted by Notre Dame's Sociology Department and supported by their History and Philosophy of Science Department. It was timely, occurring in May, a couple months before my official graduation date, and after I had finished putting together the first draft of the booklet. Unexpectedly, attending this conference weekend was vital in refining what I wanted to do with my thesis.

First of all, one of the most contentious topics during the conference was whether or not the word "measurement" described only scalable quantities, or if this word also encompassed more qualitative assessments. One scholar in particular continued to contest, against the majority opinion otherwise, that "to make everything measurement is to make nothing measurement." He insisted that other forms of knowing (besides quantification) indeed existed and were indeed important, but
could not be described as measurement. This struck a chord: some of my initial reflections on my hypotheses were aimed to answer this particular issue! In fact, my entire project rested on the assumption that he was wrong. The significance of my project was in demonstrating how important it is to understand measurement as a larger action than simply quantification.

This wasn't the first time I had heard the argument that measurement is synonymous to quantification; several people had posed this counter-argument when hearing about my initial observations on the subject. Responding to this counter-argument became the jumping-off point in my 25-page hypotheses document. I turned the “measurement = numbers” argument on its head, pointing out that there exists three (and probably more) ways to assign numbers to an object, event, or process: cardinality (amount), ordinality (order), and nominality (name). Ultimately, numbers are arbitrary tools used to signal relationships between sets. More specifically, they assign an element (an object, person, or event, etc.) as having membership within a set, which can then be compared with other sets—what I argue is the actual work of measurement.

This elaboration on the more intuitive concept of what it means to measure takes some extensive explanation, and will take even more work to effectively prove. I will need to establish a comprehensive understanding of the scholarship that exists external to this project, to position myself and my arguments among other scholars who have thought longer and harder about this subject than I have. After fully grasping this concept of defining measurement as the ever-developing crux of my thesis, I realized that a convincing and well-grounded essay proving the central argument that arose from this exploration was out of the scope of this project.

During the same conference weekend, I met with the Graduate Director of History and Philosophy of Science, who talked with me about my project and my aspirations for where it could lead. She too agreed with what my thesis advisor had been trying to get me to recognize, and what I was unable to accept until after that weekend: Now was not the time to expand the goals of the project, and an intense research paper, in addition to the work pertaining to the booklet and post-project reflection, was out of the question. The research paper would have to wait. Completing the thesis would give me a head start on where my research was headed, but trying to make it a part of the thesis would be unrealistic at best, and at worst, freeze and overwhelm my progress. The graduate director even suggested that this project might actually be the beginnings of my graduate dissertation. Later, when I talked with my thesis advisor about my realizations about the project, he confirmed her assessment: My attempts to resolve all of my questions about measurement within my undergraduate thesis was unrealistic, and even a little laughable. These two mentors of mine
seemed to have conspired together against my plans—and they probably saved me from getting
stuck and discouraged with my inability to accomplish all that I wanted to do with this project.

Reflection

"The world is a looking glass and gives back to every man the reflection of his own face."
- William Makepeace Thackeray

Although the conference weekend forced me to postpone my goal of forming a
comprehensive argument that elaborated on the definition of measurement, it wasn’t only useful
for showing me where I was bound to come up short. It also showed me that there are others
invested in my same questions about measurement, and that I had a valuable perspective towards
helping answer them. In short, the experience reassured me that I was on to something. Not only
did some of their ideas mirror mine (and then taken 12 steps further), but I also noticed that I had
picked up on some underlying features about measurement that those who spoke at the conference
had not touched. Perhaps this was because they were most concerned about the application of
measurement—particularly, measuring culture. While this still required philosophical inquiry, none
participating were very interested in asking something as abstract and foundational as, say, “Hey,
what are numbers anyway?”

As a student of pure mathematics, I’ve been trained to investigate a system (or an action, or
a definition) by identifying its rudimentary building-blocks. This perspective on the philosophy of
measurement is an advantageous one; I’m more interested in the foundations of measurement and
numbers than the practical applications of either. This puts me at a vantage point to see beyond the
mere applications of measurement, to poke and prod at finding the actual defining features, rather
than merely relying on examples of measurement. Any attempt at defining what it means to
measure merely by referring to examples of how measuring is used presupposes that we know
which examples “count” as measurement—assuming we already know the definition of the very
thing we hope to define. Although this may seem like philosophical gobbledygook to those who
regularly use measurements, these foundational questions drive how we view the application of
measurements and what “counts” as representational accuracy. At the conference, for example, the
goal was to decide how to measure culture accurately. So, clearly, the focus was on application more
so than theory. But, time and time again, the conversation ran head first into the same
 disagreements and misunderstandings surrounding an unresolved question: What does it mean to
measure in the first place? We can’t escape having a philosophy about measurement—that’s not a choice. The choice becomes whether to form judgements based on well-rounded evidence and inquiry, or arbitrarily based on opinion and previously held experience.

It should be noted that “previously held experience” is always limited experience—not necessarily as a result of ignorance nor of inexperience, and possibly as a result of just the opposite. Intense study in a particular area of knowledge vests one with the necessary methodologies and vocabulary for study in that area, in turn, defining the way one approaches a problem and what counts as a solution. Expertise necessarily creates a perspective from which some things are counted as important, others not as important, and some things irrelevant. The answer to “What does it mean to measure?” will thus be given in terms native to one’s own field, employing the backdrop gained from years of research in particular contexts that matter, ignoring contexts that do not. Thus, crucial to any results from a discussion of measurement is who is having it. My earlier question of “How many people do I need in order to talk about this?” becomes copiloted by a stickier issue: Who has a right to this conversation?

For example, if we presuppose that only disciplines which measure with numbers have the right to form a conclusion on what it means to measure, then we will get the (potentially) incomplete and self-fulfilled answer we deserve. This is not to suggest that experts of one field have no concern or regard for the perspective offered by another field, but to acknowledge that they may not even have the necessary tools to think from other perspectives in the first place. As earlier stated, the nature of expertise is to have a wealth of vocabulary and methodologies from which to frame problems and guide the expert’s thought process. And, these framings inevitably render some aspects of an issue as largely invisible—which is why the format of this creative project is so important to how questions of measurement can be answered. The project began by acknowledging and calling on several ways of knowing, and culminated in the creation of a booklet to (literally) bind these perspectives together, simultaneously presenting a multi-dimensional view of what it means to measure. There is no evil in expertise, any more than there is evil in seeing two-dimensionally in a three(-or-more)-dimensional world, as every human does. Yet, an acknowledgement of the shortcomings of any perspective allows for the humility to collaborate with others—to admit that we are all non-experts in this great universe-ity.
The following collection of essays was saddle-bound as a booklet, a copy of which was provided to every contributor. After the booklet, I include an in-depth reflection on how the arguments presented in the collection resonated with (or challenged) my own pre-project hypotheses. I should note that none of the participants were aware of my hypotheses before writing their essays, and were only provided with the four questions in the project invitation (see appendix). Thus, when reading my project reflection, note that the connections from the essays to my hypotheses are circumstantial, interpreted through my own inference as being in “response” to my thoughts.

The project reflection also became an opportunity for me to show how these essays “speak” to one another other, and I make brief notes on some of these connections throughout. Lastly, I outline some of the running themes, concerns, and observations regarding measurement that occur throughout the booklet, most of which were entirely new to me. This last section of my reflection is in the form of an outline; I merely note which scholars approached each theme and quote the portion of their essay which touches on that aspect of measurement. I encourage readers to form their own interpretational bridges between these quotations, and evaluate for themselves whether the sum of the multiple perspectives on each theme is greater than wisdom of its individual parts.
THIS IS NOT A PIPE:
ESSAYS ON MAN AS THE MEASURER OF ALL THINGS

Edited by
Rebecca L. Jackson

A Ball State University Honors College Thesis Creative Project
This Is Not a Pipe: Essays on Man as the Measurer of All Things
CONTENTS

1 Introduction  
Rebecca Jackson – Editor

3 Measurement in History  
James Connolly – History

5 Measuring a Life  
John Emert – Mathematics

6 The Distance Between  
Jennifer Blackmer – Theatre

9 Defining Literary Greatness  
Patrick Collier – English (Literature)

11 Measuring: Who Counts  
Juli Thorson – Philosophy

13 Spanning Silence  
Mai Kuba – English (Linguistics)

15 On Choosing a Metric  
Anonymous

18 STANDARDIZED BUT NOT  
Todd McKinney – English (Creative Writing)

20 Measurement in Economics  
John Horowitz – Economics

21 The Sociologist’s Tool Chest  
Melinda Messineo – Sociology
23  Measuring Influence  
   *Jason Powell - Honors Humanities (Intellectual History)*

24  Four Quartets: a Study of Metrics  
   *Brent Blackwell - English (Holocaust Studies, Literature, Rhetoric and Composition, Interdisciplinary Studies)*

30  Measuring in Computer Science  
   *Paul Gestwicki - Computer Science*

32  Measuring a Growing Thing  
   *Rebecca Jackson - Editor*
The scholars who appear in this collection each received an invitation one day, asking them to ponder the idea of measurement and create an essay response to any of the following four prompts:

1. What is the role of measurement in your field?
2. What does it mean to measure with accuracy?
3. Describe a metric within your field, or perhaps the standard "ruler" you use.
4. Describe a problem with taking good measurements, or a problematic effect of taking measurements.

Those who elected to participate were also let in on a dirty secret: They didn't actually have to directly answer any of the above. The one and only constraint was that the response had to be at least one word. Thus, the variation of responses unto itself tells its own story: Some essays stick closely to a prompt, others do not. The quickest response I received was from an economist, who, once invited to participate, promptly wrote his essay with characteristic efficiency—while I was still sitting in his office. The longest response was from an interdisciplinary professor, who answered all four questions from four different academic perspectives. Some essays include citations, many do not. Within this collection you will find personal anecdotes and observations, song lyrics, metrics, statistics, unanswered/unanswerable questions, and among all of these, the kind of deep understanding it takes decades of discipline to build.

Together, these pieces create a collage of knowledge and ways of knowing, all centered on the role of humankind as measurers. Each essay illuminates the others, surfacing new questions through the unexpected connections it makes unknowingly with its neighbors. I encourage the reader to read the essays both in and out of order. The order is intentional, and tells the story the editor saw from the connections she gathered from reading the completed work. Yet, let
the reader decide what her own path through the collection will be, and an entirely different story may arise.

This collection acts as kind of Rosetta Stone, with the “language” of each perspective on measurement adding to the understanding of the others. This allows the reader to bounce back and forth between the foreign and familiar, using her own previously acquired knowledge to piece together an understanding of unfamiliar vocabulary or calculations. It is not new or profound to state that what distinguishes the boundaries between academic areas of study are the methodologies, metrics, framings, and terminologies specific to that discipline. The lucidity of this fact, however, overshadows a broader truth: If different methods of measurement are what separate and distinguish various disciplines and ways of knowing, then measurement is also the uniting feature underlying all human pursuits towards understanding.

There has never been an age with more information (or more access to it) than the one in which we now live. There has never been more scholarship or more research set into motion to meet the demands brought on by a globalized age of infinitely complex problems—problems that are, across all cultural and disciplinary borders, shared. If we cannot avoid sharing problems, we cannot afford not to share solutions. My hope is that this collection draws attention to the commonalities between areas of study, blurs the lines between the goals and struggles of divergent fields, and creates a dialogue between different methods of measuring.

In the pages that follow, read the thoughts of scholars who took time out of their schedules, who stepped outside of their traditional research agendas, knowing full–well that this action would not be rewarded by any departmental evaluation or result in a prestigious publication. They wrote, despite knowing that their efforts would go largely unmeasured (but not, I hope, unnoticed) by their university.
For the historian, measurement is seductive. It promises to help us comprehend the past with precision and certainty. Like all seductions, there is deception involved because human experience—the thing historians seek to capture and reproduce—is immeasurable.

At the least since the nineteenth century, historians have sought a "scientific" means of recapturing the past, and the possibility of measuring social phenomena made that ambition seem realistic. The lure of producing history that explained everything with exactness flourished during the nineteenth century and has retained its appeal ever since. Henry Adams proposed that the second law of thermodynamics, a mathematical formula that tells us that the energy of the physical world will dissipate over time and disorder will follow, could also explain historical change. He measured rates of suicide, substance abuse, and other behaviors to buttress his case. Theories and methods drawn from science have continued to attract scholars. By the middle of the twentieth century, historians turned to statistics as a tool. They believed that by studying birth rates, voting totals, gross domestic product, and other empirical measures, they could make sense of the collective experiences of workers, slaves, housewives, and others, the sorts of people who did not give speeches or write memoirs. Over time, these methods grew more sophisticated, as scholars employed techniques as such regression analysis or, more recently, harnessed digital tools to make sense of "big data." Each step in this direction held out the prospect of a more comprehensive and accurate understanding of how ordinary people lived and societies evolved.

Yet the fundamental thing these historians sought remains elusive. Counting votes or tabulating intermarriage rates leaves unanswered the biggest questions: Why did people do what they did? What meanings did they assign to their behavior? How did they perceive the world around them and their
place within it? Answering these questions gets us to the essential, defining elements of human experience. But they are subjective answers, the product of interpretation rather than objective measurement.

I have been involved in a research project that has brought these truths to bear in my own work. A little over a decade ago, a colleague discovered a cache of century-old circulation records in our local public library. We have collaborated with colleagues and students to create a database that links borrowers and the books they selected over an eleven-year period. This remarkable resource documents 175,000 loans; it allows us to measure the popularity of texts, authors, and genres, to determine who borrowed what, and to gauge how often they borrowed. We can even use data visualization techniques to see relationships among books and borrowers beyond what a rank-order tabulation can tell us. This data illuminates patterns of historical reading behavior that have been largely invisible. But answering the most tantalizing question raised by this project remains just beyond our grasp. We cannot gain access to the reading experience itself. Knowing what someone read is not the same as knowing what they made of what they read. The thoughts and images they conjured up as they engaged with a text remain elusive. To comprehend those things requires using other, less measurable forms of historical evidence, and even then we can only generate imperfect, secondhand estimates of what once resided in an individual's imagination.

Measurement does not have to imply a purely objective, empirical process. After all, we sometimes speak of "taking the measure of a man" when we assess someone's character. For historians, though, it has too often meant a kind of false precision, a sense of unwarranted certainty about our predecessors and the world in which they lived. We are at our best when we measure past experience in the broadest sense, while resisting the temptation to believe we can wholly recapture what has gone before us.
The song which opens the second act of Rent asks, “How do you measure a life?” Musicals are measured by their endurance and their net profit. An actress is only as good as her next show. On the academic stage, faculty careers are too often measured by counting publications, presentations, patents, and grants. We acknowledge that quality varies, we acknowledge that impact and applicability is important, we acknowledge that collaboration and engagement is critical, and yet too often we fall back to the quantum path when we are tasked to measure a career. We reflect our values through our act of measuring: by what we measure, the means by which we measure, and the level of commitment that we dedicate to the process of measurement.

We must remember that measuring always has an impact—at times, a life-changing impact. The Rent song advises us to “measure your life in love.” The most appropriate measurements may not be quantifiable or well-defined. But perhaps it is not the measurement that is most important, but rather the act of measuring.
How on earth do you measure a work of art?

It's impossible; we shouldn't attempt it, because art is subjective, right? But we can't NOT measure art. We're compelled to do it all the time:

- "That movie was awesome."
- "One of the best books I ever read."
- "That was the worst play I've ever seen."

While it may not seem as though these observations are actual measurements, per se, they are—qualitative measurements, not quantitative—and they evaluate the experience of one work of art set against others:

- "That movie was awesome." (As opposed to other movies that are NOT as awesome.)
- "One of the best books I ever read." (As in, high above others on a scale of best to worst.)
- "That was the worst play I've ever seen." (See above, and try not to vomit a little... oops, too late.)

These measurements, while meaningful, perhaps, to those who make them, are rarely useful to anyone beyond the person who actually HAD the experience. (Think about the last time you enthusiastically recommended a movie to a friend, only to be greeted with a mild "eh" and a shoulder shrug when s/he saw it.)

Yet, these imprecise evaluations are the source of endless frustration to those of us who make art. We hate them. But they will never go away because they're easy, and for someone not well-versed in making and thinking about art, they're
the only vocabulary they have. (In addition, there are people called critics whose primary responsibility is to evaluate art, and tell us whether or not a piece of art is worthy of our time. Which is also frustrating because the critic is simply one human being who responds to a piece of art on one particular occasion, and if she happens to be in a bad mood, or recovering from a nasty case of food poisoning, then her response will most certainly be filtered through her experience, whether she wants to admit it or not. And we hate critics, whether we admit it or not. Which we won't.)

And that leads me to the next question:

*What are we actually measuring here?*

All forms of art, especially making plays and performances, are really conversations between a creator and her audience. And in any conversation, there is a GAP of understanding, the bridge of time and space that must be crossed between two individuals who are seeking to exchange meaning and truth with one another. **That to me, is the critical measurement we must try to make—measuring the length of the divide between souls.**

Sometimes it's a great distance:

- “Will there ever be peace between Muslims and Jews?”
- “What happens when we die?”
- “What is the real function of interpersonal communication in an age ruled by technology?”

Sometimes it's small:

- “I'm a lot like Tina Fey.”
- “Look at those hilarious cats in that online collection of hilarious cat videos!”

Measuring that distance *between*, no matter how large or small it may be, no matter where it is, is a way for us to get one step closer to understanding and appreciating the value of art for all people. Evaluating how large the gap and how effectively that distance was crossed, if it was, can lead us to a common vocabulary that doesn't diminish the wants and needs of the artist, but provides him with valuable information to keep making art, which is something the world can never have too much of, right?
But this is not an essay.

It's a document containing my thoughts at 1:39 PM on a Tuesday.

And it's not even a document, really. It's a collection of zeros and ones that describe the ILLUSION of a document in the form of a process that creates an image on a screen...

And I could change my mind, which is a good thing, ultimately.

But before that, I think I'll take a twenty-minute nap.
DEFINING LITERARY GREATNESS

Patrick Collier

Historically, those of us who study literature have not been centrally concerned with precise measurements. The exceptions would be the branches of the field known as bibliography and book history, where the exact number of editions issued of an eighteenth-century devotional text, or the aberrant size of pages produced by a single English printer, can become matters of import. Since the mid-twentieth century, however, literary study has been dominated not by bibliographers or textual editors but by critics, for whom the word “measure” is likely to come up only in reference to a certain problematic Shakespeare play of 1595, whose title contains that word twice.

But to say that precise measurements are not part of today’s literary scholarship is not to suggest that acts of measurement never occur. Indeed, we are tacitly measuring whenever we assign value to a literary text: that is to say, anytime we read a novel, a poem, a play, or an essay and decide whether we think the text is valuable, and if so, how valuable, and, if very valuable indeed, valuable for what purposes?

Everyone makes value judgments whenever they “consume” any kind of cultural text—arguing with our friends about the quality of a movie we just watched together is at least half of the fun of seeing the movie itself, and the top 10 or top 100 list is a major trope of American life. But as professionals whose work consists substantially of leading students through intellectual investigations of cultural texts, literary scholars are constantly evaluating texts not only as pleasurable leisure experiences but as potential grist for the classroom. This novel was fun to read on vacation last summer, but does it have the heft to bear close analysis? This poem gave me hours of geeked-out, interpretive pleasure, but will it, as the phrase goes, “teach well” to sophomores?
One peculiarity of the history of my field is that for decades we never talked about the relative quality (or relative value) of literary texts. In the mid-twentieth century, the New Criticism emphasized the painstaking close reading of a relatively small number of canonical texts. Raising questions about their value seemed beyond the point. It took the consciousness-raising movements of the late 1960s and the revolution in literary theory that followed to shake the foundations of the canon and re-introduce the heretical question of why one would read one text rather than another into the discussion. While this game-changing development began with a desire among students (and sympathetic professors) for a range of literary texts that more accurately reflected the real diversity of American life, it ultimately produced a rigorous and revolutionary theorization of literary value itself, in the work of such critics as Pierre Bourdieu, Barbara Herrnstein-Smith, and John Guillory.

Where has this left us? Less certain of the validity of our own value judgments, certainly. Aware, as well, that whether a book (or a poem) speaks to a reader depends in part upon that reader's experience—a fact differently, but no more or less, true of a professor with a Ph.D than an eager but overmatched sophomore entering her class for the first time. Aware that the literary value of any text is a process that occurs in history—a complex result of the interactions between writers, texts, readers, schools, and the many evaluative voices that social media has only seemed to multiply in recent years. When we think of a book as "great" today—and don't be deceived, we do—we are apt to be thinking pretty specifically about what, for whom, and under what conditions others might perceive it to be great—and what exactly it is great for. We've ceased, in other words, to see greatness in literary texts as simply a property of the text itself and rather as a result of a complex and unpredictable set of variables. With our local power over syllabi and examinations, we are one, but just one, of these variables.

All of this is to the good, I would say. But it means that our "measurement" of greatness in texts is anything but precise or absolute.
When thinking about measurement, a philosopher may come up with questions such as:
1. How do assumptions about measurement impact our ability to know?
2. What would it mean to measure accurately?
3. What are the norms of evaluation with which we measure an argument's worth, or the worth of a belief?
4. What are the problematic epistemological or moral results of measuring?
5. Who measures and on what authority?

In the not-too-distant past, philosophers would have been in fairly uniform agreement that one of the most important measurements is packed into the “correspondence theory of truth.” All this means is that our sentences somehow match, describe, or mirror the world. We take the measure of the world by describing it. Yet this common assumption has run into difficulties. Immanuel Kant argued that we never experience the world. Rather, we take to the world organizing principles that impose order on it. These assumptions are about space and time. Kant claimed that the world as it is (the world in itself) is beyond our ability to know. All we can know is the “phenomenal world” which is experience shaped by our organizing principles.

As influential as Kant was and is, the correspondence theory of truth remains a powerful assumption. None of us can get through a day without it. We assume that we know the world because our descriptions have taken its measure. All of us act on measuring the world in this way. There are worries, however. What we see in the world and what we find worthy of measure is the result of epistemic and moral assumptions. These underlying assumptions structure how and what we measure so that, in general, those in power do well. Traditionally the tools with which we measure have been evidence, justification, and truth, all of which are good and important. Nonetheless, traditionally those who
have measured assumed a uniformity: all measurers are rational, unemotional, privileged, and most likely male. There is only one way to measure and it is a good way to measure because there are no complicating factors, so tradition assumes.

Many feminists would argue that centering on evidence, justification and truth without taking into consideration who is measuring will produce a skewed result. This is not to say that evidence and justification are unimportant; indeed, they still are. Yet there is an asymmetry when it comes to whose measure is taken as authoritative. Some argue that those who lack privilege can often see in ways that those of privilege cannot. The material and cultural circumstance of those not in power provides a different way to describe the world, and hence, a different way to measure it. Not all measurements are acceptable. Some lack the evidence to claim that they describe something real in the world. The evidence of those on the down side of power, however, cannot be ignored if we want to approach a more complete, and perhaps a more accurate, measurement.

If we return to the questions above, we may be able to say something useful now. Assumptions about who can measure can negatively impact what can be known if those assumptions limit who can be a measurer. What counts as an accurate measurement and what you want to measure will depend on your starting assumptions about who can be a measurer, how we measure, and what is important to measure. Do you want to accurately measure medium sized physical objects? If so, the correspondence of truth will serve you well. Do you want to accurately measure who can know well? If so, the correspondence theory of truth will be of little value. If we think about who is measuring, do the norms of evaluation with which we measure an argument's worth or the worth of a belief shift? Is there an unbiased way to measure each other for how well we measure the world? How do we grant authority to someone as a good measurer? Perhaps the measure of our humanity will be the breath and depth of who counts as a measurer.
In linguistics, we quantify many aspects of speech. How often do speakers say certain things, how loudly, at what pitch? However, it may be even more important to measure the absence of speech: when do participants in spoken interaction fall silent, and for how long?

Conversational pauses matter. Pauses can influence how fluent a second language speaker seems to others (Watanabe & Rose); when pauses are unexpectedly short or long, they can contribute to perceptions of rudeness or lack of engagement (Rimondini 2010). Since norms for pause duration vary across cultures (Edwards 1993), these perceptions can lead to inaccurate negative impressions in cross-cultural interaction.

Measuring the length of a pause is not a simple matter. For one thing, it is not physiologically possible to talk continuously, since we need to breathe; moments of silence associated with respiration should perhaps not be counted in the same way as silent pauses that have a communicative intent. Another consideration is that, when a pause occurs between turns, the researcher cannot know whether to attribute the pause to the person who just yielded the floor, or to the person who is about to speak (Kendall 2009).

The matter of duration itself brings more challenges: how long must a silence be before we count it as a pause? If a silence is shorter than 200 milliseconds, it is difficult to distinguish from certain consonant sounds which inherently involve a short silence (Kendall 2009). (This is the case with the initial sounds in the words pat, tat, and cat, although listeners do not perceive silence in those initial sounds, of course.) In many studies, silences at least 270 milliseconds in length are counted as pauses; the upper threshold may be as long as 5 seconds (Kendall 2009).
Pause duration can be measured with precision from a computer display; however, in many studies, especially in the earlier days of conversation analysis, stopwatch timing or even perceptual judgements were used (O'Connell & Kowal 2012). These low-tech methods may seem shockingly imprecise; and yet, depending on the researcher's goals, they have some benefits as well. In interaction, the objectively measurable duration of a speaker's pause may be less important than the listener's subjective impression of it, and that perception depends on factors such as the speaker's speech rate (Edwards 1993). Therefore, "to approximate the likely perceptions of interactants, some researchers quantify pauses as the number of beats of silence, based on speaker's preceding speech rate" (Edwards 1993).

It is both important and challenging to measure conversational pauses. When studying language, we have to attend to not just the words themselves, but the spaces between words. This enterprise challenges us to grapple with the potential conflict between objectivity and subjectivity.

In mathematics there are a vast number of metrics used to measure a wide variety of objects. Perhaps more interesting is that a single object can be measured by many metrics, which could give quite different impressions to “users” of mathematics who do not take the time to understand these metrics. Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking that their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind. Though properly chosen metrics can do just that for us, these “users” often don’t take the time and energy to learn what these metrics are (and are not) and can often make gross errors in their data analysis.

As an example of how metrics can differ, consider a sequence

\[ A = (1, 3, 4, 8, 1) \]

and consider the problem of determining which of the following sequences B, C, and D most closely resembles A. That is, we are looking to measure how close the following sequences are to A. For a meaningful interpretation we may consider the components of A, called \( a_i \), to be the stock price of a company \( i \) years ago. Hence last year’s stock price was \( a_1 = 1 \), and the stock price 2 years ago was \( a_2 = 3 \), and so on. With this in mind, how should we determine which of the following sequences of stock prices for other stocks are similar or dissimilar?

\[ B = (2, 2, 2, 2, 2) \quad C = (10, 2, 4, 8, 1) \quad D = (1, 3, 4, 8, 10) \]

A common metric is to measure the “distance” between sequences

\[ X = (x_1, x_2, ..., x_5) \text{ and } Y = (y_1, y_2, ..., y_5) \]

is

\[ d'_i(X, Y) = [(x_1 - y_1)^2 + ... + (x_5 - y_5)^2]^{1/2}. \]
Using this we see that $d(A, B) = 6.56\ldots$, $d(A, C) = 9$, and $d(A, D) = 9$, and so $A$ is closest to $B$, while $C$ and $D$ are equally close to $A$.

Another useful metric that will measure the “distance” between sequences $X = (x_1, x_2, \ldots, x_s)$ and $Y = (y_1, y_2, \ldots, y_s)$ is

$$
d_l(X, Y) = |x_1 - y_1|/2 + |x_2 - y_2|/4 + \ldots + |x_s - y_s|/32.
$$

Using this we see that $d_l(A, B) = 1.40625$, $d_l(A, C) = 4.5$, and $d_l(A, D) = 0.28125$, and so $A$ is closest to $D$.

Another useful metric that will measure the “distance” between sequences $X = (x_1, x_2, \ldots, x_s)$ and $Y = (y_1, y_2, \ldots, y_s)$ is

$$
d_\infty(X, Y) = \max(|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_s - y_s|).
$$

Using this we see that $d_\infty(A, B) = 6$, $d_\infty(A, C) = 9$, and $d_\infty(A, D) = 9$, and so $A$ is closest to $B$, while $C$ and $D$ are equally close to $A$.

Inspection of the formulas shows that $d_l$ treats all coordinates equally, meaning that if you were to, say, switch two coordinates, like the first and third in each of $A$, $B$, $C$, and $D$, then the measurements would remain the same. This is also true for $d_\infty$, though unlike $d_l$, the metric $d_\infty$ only cares about the largest difference in stock prices over the years whereas $d_l$ takes into account differences for all years. However, $d_\infty$ utilizes a weighted sum and so the first coordinate is weighted the most, with progressively lower weights for each following coordinate. This is particularly appropriate when for modelling stock price futures since it is often the most recent stock price that is far more important than those of the distant past. Many phenomena follow such a model. For example, the weather tomorrow is much more tied to the weather today than it is tied to the weather 30 days ago.

Just as the different metrics used above can be used or avoided depending on what the metric is supposed to measure (with respect to what aspect the scientist is studying), the same is true in image compression metrics. When comparing two images, say, a digital photo with a jpeg compressed approximation, there is a standard metric that is used for many settings and so often gets applied to images as well. This is just a variation on the metric $d_l$ above, a standard metric used in statistics as a default since it is well-known and
simple to apply. This metric, just like $d_r$ above, treats all coordinates the same. When desiring a metric to measure how similar to the human eye two images are, however, this is not the best metric to use. Instead, it would be best to use some metric which weights different parts of the picture differently. The reason is that for the human eye to distinguish or recognize an image, the boundaries between objects in the picture are far more important than what is going on where there is little or no variation in the color. For example, when looking at a red square on a white background, the eye spends little or no time focusing on the large white regions outside the square nor on the large red region inside the square. It instead focuses on the boundary to quickly recognize both “red” and “square,” and then the brain can easily store that information and match it with memories of stored pictures already seen. The useful information in the picture is in the boundaries, and so the metric should take this into account. These metrics are more complicated to use, but they are much better for this situation.

Knowing of the wide range of metrics that can be used in a vast number of applications, we should all keep in mind whether or not metrics were chosen for being most suitable to the aspects we wish to understand or whether they were chosen as a default.
This week, my son finished his first round of ISTEP tests. As a third grader, he only has nine more years of these tests to go. While the tests measure the schools' work, it's hard for me not to stress about these tests, especially as students have ID numbers behind which all test scores will forever follow. How will teachers and administrators judge him? This year's test is a pilot test being given to all the schools. Testing companies are even developing tests, I've heard on the radio, to record the speed with which a student answers a question, but also his or her hesitation. What if a student was having a bad day? Like her dog ran away? Like he was out of his favorite juice that morning? Today, I listened to some music (Neko Case, the Wynton Marsalis Septet, Bob Dylan, etc.) while working on a table I'm building, I had to unscrew one board from two pieces, and then reconnect it, so that the bubble in the level was between the lines. The bubble shifted a little when I moved it from workbench to floor. I cut another board, sanded it down, and fastened it to the end pieces. Looking at the boards for the tabletop, I may have cut them too short. I am taking too long to respond to some student papers. I will respond with positive comments, critical suggestions, some questions, and some grammatical corrections. I might say, This idea has some real energy to it. Keep exploring it. That won't be a lie, but I can't promise that further exploration will bring about more light. How does one measure development? The more I've lived, the less I know. Or another way to put it: the more I do, the more aware of my own ignorance I have become. We are each on our own journey. Do your best. It's convenient that one Starbucks is like another. The coffee, I know from my past experience, will most likely be satisfying but will I remember it? Sometimes, I listen to a song over and over, like Neko Case's "Nearly Midnight, Honolulu," a song like no other that I've ever heard. You'll have to Google it. If a student writer can level a truth, an insight, an emotion as clearly or powerfully or as poignantly or as ________ as Neko Case does in this song—and I've seen it done and I can recall a number of essays and poems and stories students have written that do
this—then I celebrate it. I try to let the writer know what they’ve done. Often, though, it takes at least several drafts for a writer to capture it. Some need more, some need less. Elizabeth Bishop needed seventeen drafts to get to the final version of “One Art,” a great poem. It’s a villanelle in iambic pentameter with some variation from a strict meter. I have learned how to measure a poem, etc. by reading a lot of poems, etc. The joy of a good poem, etc. is that it’s not like any other, or that it’s different or new in some way even if it conjures up another poem, style, story. Generally, one does not write inside a closed room. The instructions for the table I’m building I found online. Like you can find the rules for a villanelle online. Ultimately, I learn by doing. The last one I built, most of my friends have admired it. Is writing much different? We want pieces that hold up, that hold some beauty or insight or _______. Pieces that seem to be made by an individual human spirit, one who knows (or strives to know?) something about the daily struggle that she or he has put into words memorable and sublime because, despite being made of an ancient alphabet and all-too familiar themes (grief, injustice, etc.), the words are for their ideas and order ringing with that sense of being well crafted and somehow personal. And that seems worth a huge smiley face sticker and a standing ovation.
Economics is the study of how limited resources are used to satisfy the highest valued of people's unlimited wants. In other words, it is the study of choice. Choices are usually improved when people have better information about the world. This is the purpose of measurement: To help us understand the world and make better choices, because you can't always get what you want. Measurement helps us understand the trade-offs that we must make.

One example of measurement in economics is the poverty benchmark. Poverty is normally considered as not having sufficient resources. However, there are several issues. Should there be different standards in different countries? Do people in the Congo where poverty is defined as a person living on less than $2 a day need less resources than a single person in the United States who needs about $32 a day? Should the standard vary by family size? If a person in the United States lives in a family of four, they need at least $16 dollars a day rather than $32. Should all income be included? In the United States, not all income is included. For example, in-kind transfers such as SNAP (Food Stamps) are not included nor are tax transfers such as the EITC. How should people's incomes be measured? Many are asked to report their income, and studies have found that many do not report all they receive in transfer payments. And how much they report varies by program. Does the measure change people's behavior? If money is given to reduce poverty and you can get more help if you are a single parent, will some people choose not to get married in order to retain benefits?
The question of measurement is central to the field of sociology. We run on a continuum between the points "The world is socially constructed and therefore accurate measurement is impossible," to "The world is knowable and quantifiable and through this quantification predictions can be made."

As a result we have two broad ways of conceptualizing measurement: qualitative and quantitative. On some campuses, these conceptualizations divide the members into two camps and sometimes these camps are at odds with one another. There is often a sense that the quantitative work is more highly valued than the qualitative. However, on other campuses they are viewed as "tools in a tool chest" and whether or not a researcher uses a specific tool depends on the research question being asked. Some questions are better suited for different approaches. On our campus, we have lots of faculty that use both measurement approaches, depending on the project. Since becoming an expert takes a great deal of time and investment, we tend to end up developing skills in specific areas, but the orientation is toward inclusion.

One challenge that sociology faces in measurement is that we focus on and study groups and large social structures and institutions. However, in order to study a group, you have to talk to individuals. Take for example, families. Let's say I wanted to see how families experience an economic recession. I can look at collective data about their financial well-being and assets, etc. but at some point I am going to need to talk to individuals to get information. This dualism between the individual and the group is a challenge to the field.

Similarly, we often do research that has qualitative and quantitative elements. For example, I look at representations of minorities in advertisements. How this data works is that I look at thousands of advertisements and code them on a variety of qualitative criteria. Is the character prominent? Attractive?
Respected? I then quantify those findings and run regression analyses on them based on various hypotheses about the media. This process of taking a qualitative assessment and making it quantitative is potentially controversial in the field. There are many mechanisms that we have to put in place to make sure that we do not introduce bias into the process at any stage. These efforts include working toward inter-coder and intra-coder reliability. In these contexts the question of measurement becomes a test of agreement between coders (inter) and even with the coder themselves (intra).
There are many quantitative historians, those who look at the numbers of whatever they are studying: crime, number of children per family, economic costs, etc. As an intellectual historian, in the words of my advisor in graduate school, “We don’t fucking count anything.” So measurement becomes rather murky. How does one measure ideas? Some historians count the number of copies of a certain magazine or book that were purchased, to give some type of clue as to an idea’s intellectual capital. But purchasing such material does not guarantee that the ideas spread as the author intended. Maybe part of the author’s argument had some influence, or maybe people reacted against it for particular reasons. There is a history of “reader response” but I think it’s kind of boring, and it doesn’t seem to get at what it is about an idea that endures.

As a historian of existentialism, I try to see trends in certain artists and intellectuals that may have existential leanings, but since existentialism is about self-realization, how does one measure that? Sometimes it comes down to an author’s innocent statement in a letter or another piece of literature that lets me know that he or she arrived at a certain conclusion. I then look at subsequent works of that author and see if the pattern plays out. That “playing out,” to a degree, is a measurement of influence, a symbol that an idea has taken root. So, for me, measurement is not about quantity. It is a qualitative assessment as to how an idea works itself out. In a sense, I try and measure a trajectory of an idea, see how it grows and flourishes or wanes and dies. I have never thought of this process in terms of measurement, but I think it is. I measure the enlargement or diminishment of ideas.
Response as a Holocaust Professor:

*What is the role of measurement in your field?*

Measurement plays a significant role in the field of Holocaust Studies. Aside from the usual necessity for historical accuracy, Holocaust Studies has a rarefied sense of precision with respect to numbers, dates, and the likes because of professional Holocaust Deniers. While no mainstream Holocaust Denier (i.e. one who is not a certifiable, raving lunatic) rejects the premise that the Holocaust did in fact occur, they are very crafty in their denials of the minutiae. For example, one famous Denier claims that the 6 million figure is too high. He estimates the number to be closer to 1.5 million (as if a smaller number somehow lessens the atrocity or the gravity). In fact, they use discrepancies in the reporting of dates and figures on the parts of both scholars and survivors as indicative of a conspiracy to obfuscate the true facts of the event. Because two survivors’ accounts differ with respect to what most would call objective or measurable data (dates, locations, numerical figures, names, etc.), they then declare that all testimony is suspect. Of course, this denies the historical convergence of truth that underpins History as a respected and quite established field of inquiry, but Deniers care little for agreement, only disagreement. Ignoring the thousands upon thousands of eye-witness testimonies, legal documents, newspaper reports, photographs, and physical artifacts, they instead focus on the fact that there are small holes in the story (as there are in all historical research). Just because we cannot establish with certainty that Ernst Kaltenbrunner, chief of the Reich Main Security Office of the SS, ever visited Mauthausen concentration camp does not lessen his involvement with, or Mauthausen’s role in, the murder of 6 million of Europe’s Jews.

Another odd metric found in Holocaust Studies concerns proximity. Since this event has almost completely transformed from recent event to history, this
metric must be common in other areas like this one. But in Holocaust Studies not only is it the proximity to the event that tends to carry more weight, there is also a stark divide between those who experienced the event and those who study it. Ellie Wiesel is the most famous proponent of this division, believing that the ones who did not witness or experience the harrowing events have no right to an opinion about it. This is the call to silence he proposes—a call to not just avoid judgment (which makes perfect sense), but a call to silence about the entire event.

Coupled with his insistence that the Holocaust is a particularly Jewish event, he has proscribed a few metrics in our field that we as scholars must contend with if we want our work to be counted as legitimate by non-scholars (yet another metric not common in many fields). As a non-German gentile, my work is already suspect from his point of view and many others who agree with him. As a field, we debate on the validity of the objective position itself, as much as anyone can be in one. And to my knowledge, this is certainly one of the few fields where someone with an axe to grind or personal grudge to bear, in fact, carries more weight than those of us who have none.

Response as a Professor of Rhetoric and Composition:

*What does it mean to measure with accuracy?*

Measuring student writing is a notoriously tricky thing. For decades, composition instructors have struggled to assign an objective score to something that is essentially subjective. So the question of accuracy is suspect from the start. How does one measure writing accurately? We have taken solace in the reality that academic writing is significantly more measurable than something like creative writing (even though I disagree), but that provides little comfort to undergraduate writers.

Simply put, accuracy is something that is gained over time. It is a skill that one can measure only over multiple attempts. The more papers someone grades, the more accurate the grading becomes. Over time, we can measure student writing in aggregates. We can, with experience, support a claim that says most 'C' writing lacks certain features, or most 'B' writers demonstrate a similar level of mastery of rhetorical concepts like context or audience. With more experience, we can even refine such measurements according to pluses and minuses (i.e. many B+ papers exhibit a fair amount of in-depth critique). While such claims are inductively sufficient, they are not necessary. There will be at
least one B+ paper (in the history of every B+ paper) that does not demonstrate a fair amount of in-depth critique, but such papers are exceptions, not norms.

A more difficult issue in accurately measuring student writing lies not in the rubrics chosen for formal evaluation, as again most writing professors agree on the short list of grading rubric items (thesis, audience, context, grammar, etc.), but rather in the weight assigned each of these items. So while every writing professor agrees that grammar is an important component of good academic writing, they do not agree on the weight that it should carry in a paper, which can lead to larger deviations in grading. In our department at Ball State, for example, we collectively agree that thesis should hold the most weight of any category of evaluation. The problem comes in accurately specifying what "most" means mathematically. Writing professors are allowed the freedom to interpret "most" freely from as little as 51% of a paper’s grade to 99% of a paper’s grade. These kinds of interpretations of the rubrics often come to a head with respect to grammar. Some professors (especially outside of the English dept.) do not agree that thesis should count the most in an academic paper. They would prefer grammar carry much more weight, so they adjust their rubrics according to their own sense of "good" writing.

Response as a Literature Professor:
Describe a metric within your field, or perhaps the standard “ruler” you use.

As a field, literary studies is guided by and defined through its use of metrics. Metrics guide departmental organization, course offerings, degree requirements, and scholarly production. Organizationally, all departments structure themselves along both historical periods (c.f. Medieval, 18th Century, Modernism, etc.) and generic divisions (literature, linguistics, rhetoric and composition, etc.) first and foremost. After these divisions, they are sub-organized by field (c.f. poetry, fiction, drama, multimedia literacies, phonology, etc.). And while no one would like to admit it, each department self-segregates with respect to these metrics, usually based on the popularity and productivity of the professors who teach in these areas. In this way, such exigencies determine course offerings and the like. These are the first two metrics that shape departments.

Within my own field (Literature), there are further metrics brought to bear upon our material of study in addition to historical area and genre mentioned above. A third metric fundamental to literary studies is the metric of literary theory. There are literally dozens of acceptable literary theories that act as lenses
through which to focus a piece of literature (Marxism, structuralism, Psycho-
analytic, biographical, deconstruction, etc.). These lenses determine things like
the accepted methods of analysis used upon the material as well as even more
ephemeral things like style. And these metrics can and often do come into
conflict with each other as well: some being widely accepted by scholars, some
highly contested.

There also exists in the field a fourth metric defined by the side of the
Atlantic on which one's field rests. Historically, English literature and American
literature have stood at odds with each other in departmental priority, policy, and
direction. As the British Literature advocates see things, the very best American
literature (usually Melville) comes close to mediocre British literature. A funny
side note is Henry James, an expatriate writer. His greatness is uncontested by
both areas, and they both lay claim to him fiercely. Like James, there is also
Shakespeare, a fifth metric.

Standing as a sort of overarching, umbrella metric is Shakespeare. Almost
by definition, the field of literary studies recognizes the brilliance, importance,
and significance of Shakespeare. He is the ruler every literary professor uses to
one extent or another: some consciously, others subconsciously. Some willingly,
others unwillingly. It could be said that there would be literature departments
had Shakespeare never existed, but they would not have the influence and
significance they do today.

A sixth metric in literary studies is a sort of dilution of all the rest into daily
practice, canonicity; what counts as great literature from a given period, and
who gets anthologized and read by students as representatives of those periods.
Like it or not, even historical areas such as medieval or the renaissance (smaller
areas than say, modernism) fall victim to questions of canonicity because there
is only so much space and only so much time. For example, there are roughly
30,000 extant lines of Old English. Period. So even in a specialized course
on Old English for undergraduates, such questions of who and what to read
become quite important as no undergrad can be expected to reasonably process
30,000 lines of text. Such decisions are based on both the canonic parameters of
the field in general and the particular interests of the professor. This, of course,
assumes that such a course is even “counted” as literature at all since many
smaller departments do not even have a specialist on Old English, nor do they
offer undergraduate courses on it, which then skews the position of Old English
as a sustainable field of inquiry in literature departments altogether. All agree
that *Beowulf* is central to British literary studies, but if very little class time is
spent on it (if at all), does it at some point cease to be, in fact, literature? Metrics can have this effect in literary studies.

The last metric is perhaps the least important from a teaching standpoint, but crucial from a publishing standpoint: topicality, or better, trendiness. In any given field, there are the trending authors/works of the moment—the hot figures at any time whom everyone who is anyone in the field should have something intelligent to write and talk about. Toni Morison had her heyday in the early 90s. In the 70s, Poe dominated Nineteenth Century American Fiction discussions. In Old Norse, everyone who was anyone wrote about Hrafnkell’s Saga in the 1980s. The trends do creep into classrooms, but generally only as special topics courses.

While this may seem like a tertiary metric at best, it does determine the field to a large extent. For example, what to do with a marginal figure? What to do with someone like James Hogg, a very important author in his day, respected by all. It has been decades since anyone in Nineteenth Century Fiction (let alone literary studies in general) has had any reason to return to his works on any noticeable scale. They just seem to have fallen by the way side, a voice that simply doesn’t speak to the 21st century at all. Will Hogg cease to be counted as literature? Topicality can literally delete or insert acceptable figures into the field. A study of Hogg, however brilliant, will not find its way into PMLA. He has a difficult enough time making Nineteenth Century Literature. So topicality does act as a significant metric in literary studies.

In this way, metrics literally proscribe the field itself, with each professional tangling with at least six metrics every time they plan a syllabus or submit an essay for publication. Metrics prescribe what can be written, what can be taught, what can be said, and how. Successful literature professors navigate these waters well. Less successful ones often get mired in one or more of these kelp beds.

Response as an Interdisciplinarian:
Describe a problem with taking good measurements, or a problematic effect of taking good measurements.

Interdisciplinary scholars, by definition, are bound by a number of metrics equal to or greater than the number of fields they integrate into their work, so taking good measurements is relative to the field that evaluates such measurements. As a Holocaust Studies professor, I integrate literary study,
political science, history, psychology, rhetoric, and religious studies into my professional work. What is maddening, however, is that these are just the fields I use. It is possible and quite common for another Holocaust Studies professor to work completely within Sociology, Anthropology, and Communication Studies. Another might work with Art, Music, and Theater. Yet, all three would be called Holocaust scholars, even though there is, in fact, no point of connection between their work save the broad topic of the Nazi Holocaust. They simply have no metrics in common at all. And all of this disregards specialties within the field. Some study the Nazis, some the Jews. Some focus on Romania or Hungary, while others concentrate on France or Great Britain. Some of us only read and study materials in German, while others only use Polish or Russian. It is possible to be a Holocaust Scholar and not even read a word of German or Yiddish, if the focus is on Holocaust commemoration and memorialization in America, for example.

From the point of view of traditional disciplinary studies, Interdisciplinary Studies is quite impossible, since the very idea of commonly held metrics prevents any real engagement with the material on a real, substantive level. If disciplinarity is defined by the metric it uses to engage the world, then real world phenomena that exist in multiple states at the same time have to be reduced, diluted, and filtered. Interdisciplinary Studies denies this singular outlook on studying the world. To truly understand and engage any phenomenon, it must be tackled on its own terms, which are never singular. Rather than trying to reduce or encompass, Interdisciplinary Studies tries to open and expand. Experience is a plenum, not a monad.
The discrete and formal measurements of Computer Science reflect its foundations in mathematics and electrical engineering. Computer programs are composed of unambiguously parsed statements and expressions, which are compiled into atomic hardware directives. Information is transmitted over networks by measuring electronic pulses and radio waves, interpreting these as bits and bytes. Computational theory tells us that all computable problems belong to a single space-time hierarchy and that some problems cannot be computed at all. This leads us to the biggest question of Computer Science—the question of whether \( P = NP \)—which is fascinating not only because we do not know the answer, but because we also do not know if it is answerable.

These discretely measurable aspects of the discipline complement a more humanistic side. Creativity and ingenuity are required in determining what line of code to write, what proof technique to use, or what software architecture to follow. First attempts almost always fail, and so we learn to approach problems iteratively, using formal and informal heuristics to determine whether we are moving closer to a desirable solution. That is, we recognize that doing Computer Science is a learning process, that building a system is the process of learning how to build the system. Unlike in engineering and architecture, where models are built for others to manufacture, the computer scientist deals directly with the elements that compose the artifact. This creates a rapid feedback loop for the reflective practitioner, giving rise to the model of expert computer scientist as master craftsman.

Building any nontrivial software system requires a multidisciplinary collaboration, where the tools of the Computer Scientist are combined with approaches from the social sciences, humanities, and the arts. Research consistently shows that teams with strong communication practices produce better results; how to actually conduct these measurements is a perennial research
question. Martin Fowler—a recognized industry leader software development—has gone so far as to claim that productivity cannot be measured at all, and that any attempted measurements will only produce misinformation [1].

My own scholarship involves building original educational video games with multidisciplinary teams of undergraduates. When interviewed about what most contributed to their successes, teams inevitably split their answers between the measurable and the non-measurable. They find that formal processes and communication patterns, quantitative feedback, and technical structure give them the confidence to build and revise complex software systems. At the same time, they recognize that empathy has the most dramatic effect on both team productivity and the quality of the product. These two sides, the scientific and the humanistic, are not opposed to each other: they form a synergy that can propel a team to success or, in their absence, drag a team to a near standstill.

This harmonious balance of the measurable and the unmeasurable characterizes Computer Science. Alistair Cockburn captures this paradox in his cooperative game principle in which he defines software development as a cooperative game of invention and communication [2]. Both kinds of moves are necessary for the team to create a solution of value to stakeholders and, thereby, to win the game. A winning team must make measurable and unmeasurable moves, creating technical and artistic artifacts while continuing to build empathy, all the while reflecting on how they are learning through this process.

My interest in measurement developed over the last couple years or so, as I was toying with taking methods I was learning in one class and seeing if they helped me better understand problems in another class. Since both of my areas of study (mathematics and creative writing) can be boiled down to the skill of allegorical reasoning—“Hey, this thing is like this other thing”—I tend to think about problems in terms of analogies. One example is the problem of measuring educational progress. The main reason I’m at Ball State is my love of liberal arts education alongside a rigorous scientific education, and we have a terrific honors college here which embraces the symbiotic relationship between the two.

“Honors doesn’t mean harder,” I tell incoming freshmen when I speak at summer orientation sessions, “honors means different. It means small discussion-based classes, being graded on your ability to communicate well-formed judgments in conversation and in writing, and professors who won’t simply tell you what to think, but will ask you what you think. Most of all, you’ll be challenged to know why you think what you think.”

Ultimately, I realized that the difference between the honors experience and the regular undergraduate curriculum was how the students were measured: Honors students were evaluated on their ability to communicate strong arguments rather than their ability to effectively memorize the right answers to test questions. But why was this important? Why was I so convinced that asking students to write papers and have conversations and craft well-formed ideas was a “better” way to measure progress, especially considering that these assignments seemed harder for professors to objectively quantify?

I finally found my answer in a class on chaos theory and fractals, during a presentation on how natural objects such as plants, rivers, and coastlines can
be best measured using fractal geometry. In measuring the length of a coastline, adding up a series of straight lines will always cut off the winding intricacies we wish we could better approximate. To solve this, mathematicians take an approximation of the whole coastline, shrink it down, then use this miniature as the new “ruler.” Because nature likes to repeat herself, we can get a much better approximation of the length of the whole by adding up tiny miniatures. This is similar to measuring the parameter of a fern’s silhouette by using a frond as your ruler. That’s the challenge, I thought. How to measure a growing thing.

It seems to me that standardized tests and fill-in-the-blank answers act as a straight ruler, a poor approximation for a much more intricate beast. When students graduate, we would like to think they have become better thinkers, better citizens, more innovative, more informed and capable; I’m not sure any educators hope students will simply be very excellent at memorizing many things for a short time and writing them on test sheets. A four-page paper recording well-developed thoughts on a single text offers a better approximation of a well-developed student, and a single class conversation can be thought of as a miniature nugget of growth in the journey towards becoming a collaborator with others. Much like how a naturally fractal object can be better approximated in its entirety by adding up tiny miniatures of itself, measurements in education ideally should reflect the larger scope of what we hope to achieve—the fern is best measured in fronds.
<table>
<thead>
<tr>
<th></th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Jim Connolly</td>
</tr>
<tr>
<td>5</td>
<td>John Emert</td>
</tr>
<tr>
<td>6</td>
<td>Jen Blackmer</td>
</tr>
<tr>
<td>9</td>
<td>Pat Collier</td>
</tr>
<tr>
<td>11</td>
<td>Juli Thorson</td>
</tr>
<tr>
<td>13</td>
<td>Mai Kuha</td>
</tr>
<tr>
<td>15</td>
<td>Anonymous</td>
</tr>
<tr>
<td>18</td>
<td>Todd McKinney</td>
</tr>
<tr>
<td>20</td>
<td>John Horowitz</td>
</tr>
<tr>
<td>21</td>
<td>Melinda Messineo</td>
</tr>
<tr>
<td>23</td>
<td>Jason Powell</td>
</tr>
<tr>
<td>24</td>
<td>Brent Blackwell</td>
</tr>
<tr>
<td>30</td>
<td>Paul Gestwicki</td>
</tr>
<tr>
<td>32</td>
<td>Rebecca Jackson</td>
</tr>
</tbody>
</table>
Post-project Analysis
Reexamination of My Pre-project Hypotheses on What It Means to Measure

At the start of this project, I collected together all of the observations about measurement that I had scribbled down during the past couple years. It was important to me that I record them before I saw any of the contributors' essays, so that I could later revisit these ideas and write about what I had learned since creating the booklet. I then distilled and organized these observations into seven major hypotheses about measurement, most of which have two components, parts (a) and (b).

In the following discussion, I reflect on how the arguments presented in the collection of essays resonated with (or challenged) these pre-project hypotheses. I should note that none of the participants were aware of my hypotheses before writing their essays, so any connections I make between their thoughts and mine are inferred. This analysis also became an opportunity for me to show how these essays “speak” to one another other, and I make brief notes on some of these connections throughout. Lastly, at the end of this discussion I outline some of the running themes, concerns, and observations regarding measurement that occur throughout the booklet, most of which were entirely new to me.

Hypothesis 1: Measurement as Perspective
Without measurement, there can exist no perspective, though any perspective has its limitations.

a. In order to have any kind of perspective, measurements must be taken from some vantage point.
b. Any single vantage point will inherently be incomplete or only approximative, and thus the resulting perspective will inherit this error.

Philosopher Juli Thorson elaborates on this idea in her comments about the "correspondence theory of truth," which assumes “...that our sentences somehow match, describe, or mirror the world. We take the measure of the world by describing it." She explains that Immanuel Kant denied this theory, instead arguing that “the world as it is (the world in itself) is beyond our ability to know. All we can know is the 'phenomenal world' which is experience shaped by our organizing principles."
She then goes on to admit that, although she seems to agree with Kant about the correspondence theory of truth, “None of us can get through the day without it.” How could we even function in our daily lives if we truly believed that the world we observe and seem to experience is not the reality? Indeed, “We assume that we know the world because our descriptions have taken its measure. All of us act on measuring the world in this way.” Measuring is inevitable in life, even if we truly believe these measurements to be mere approximations that reference reality. In the words of the French philosophers Deleuze and Guattari, “It’s nice to talk like everyone else, to say that the sun rises, when everybody knows it’s only a manner of speaking.”

This acknowledgement of error (and yet acceptance of it as inevitable) is a fundamental and necessary concession required in the act of measuring—one that I explore in further detail in Hypothesis 5.

I might add that there is still further translation (and thus potential for further error or “distance” from the reality) involved in taking human perception and turning what is observed or sensed into a verbal description. Thus, Thorson’s notion of description as measurement seems to suggest that language is an approximation of an approximation: The mind’s perception of reality is already an approximation, and the description via language is a further approximation of that perception.

Perhaps I take her use of the word “description” too literally, and what she refers to is also metaphorical: Our eyes “describe” the world to us, by recognizing certain shapes, translating light frequencies into color, giving us the illusion of three-dimensional distance, etc. Our minds also “describe” the world through pre-conceived notions of what it expects to encounter, which has potential to create inaccuracies in our perception. This is evidenced in how beginning art classes must train students to really look at an object or person, abstracted from what they think an apple or a hand should look like.

Whether or not Thorson meant description only in the literal sense, it is interesting to note how words such as “describe” or “define” can refer to both shape and language. Perhaps the act of having certain shapes and frameworks for understanding the world is a way of putting it into representative symbols—an abstract “wordage” for the mind. And, likewise, perhaps the act of describing the world in words is a way of setting some things apart, collecting some things together, and along the way, shaping it in our minds.

---

Thorson also approaches the idea of inevitable error in another way: the perspective of the measurer is often an incomplete one. She uses her brief discussion of the correspondence theory of truth to launch into a brand new idea I had never considered: *Who gets to be a measurer, and how do we decide?* She points out that those who get to be the measurers, who decide what is worth measuring and what is not, as well as which methods of measuring are reliable and worthy, are traditionally "rational, unemotional, privileged, and most likely male." I might add that these, in turn, are the describers of society, who choose the terms by which our culture is shaped, recognized, perceived.

Literary scholar **Brent Blackwell** notes (in his third response) that literary critical theories form a "lens" (or perspective) from which literature is analyzed. He refers to these lenses as "metrics"—as ways of measuring the value and interpreting the significance of texts. Often, he observes, these metrics come into conflict with one another. Like Thorson, Blackwell also notes that some metrics are privileged over others: "some being widely accepted by scholars, others highly contested."

Throughout his essay, Blackwell subtly builds a case for why multiple metrics are necessary for truly understanding complex human problems and world events. In the conclusion to his fourth response, he takes a much bolder stab at why the privileging of certain perspectives over others can not only lead to inaccuracies, but even undermine the pursuit of knowledge: "From the point of view of traditional disciplinary studies, Interdisciplinary Studies [which uses multiple metrics, none of which need be the same ones used by the colleague next door studying the same subject] is quite impossible, since the very idea of commonly held metrics prevents any real engagement with the material on a real, substantive level. If disciplinarity is defined by the metric it uses to engage the world, then real world phenomena that exist in multiple states at the same time have to be reduced, diluted, filtered." After carefully framing the problems with studying multifaceted phenomena from a single perspective, he presents the natural solution: "Interdisciplinarity Studies denies this singular outlook on studying the world. To truly understand and engage any phenomenon, it must be tackled on its own terms, which are never singular."

The use of multiple metrics at once (and the frustration with this action being considered less than legitimate) would seem quite familiar to Thorson. She (as a feminist philosopher) is likely more than familiar with *écriture féminine* and other feminist philosophies that embrace the idea of multiplicity as opposed to a single, linear path towards truth.
Blackwell’s final comments also contain a subtle “easter egg” for a scientific audience: His mention of “multiple states” brings to mind quantum mechanics, along with the many measurement problems therein. Since much of his essay focuses on the historical and literary sides of things, this inclusion of the scientific perspective could act as a signal that he writes to a widely multidisciplinary audience. Or, this use of scientific language could be more unconscious, a by-product of thinking about this issue in terms of multiple perspectives at once.

Computer Scientist Paul Gestwicki takes a more optimistic outlook on this problematic aspect of measuring. His essay celebrates the necessity of multiple perspectives, noting the shortcomings of using any singular approach. He references Alistair Cockburn’s “cooperative game principle,” which defines software development as a “cooperative game of invention and communication.” Gestwicki suggests that in order to “win the game” (create a valuable software solution), both “measurable” and “unmeasurable” moves are necessary. He also refers to these moves respectively as “scientific” and “humanistic” moves, or as “technical” and “creative” moves. This reflects an interestingly assumed association: measurable with “scientific” and “technical,” and unmeasurable with “humanistic” and “creative.”

This mental grouping is not uncommon in my own experience talking with scientific and creative thinkers. Although, if pinned down, most members of either camp would probably admit that the distinction between those worlds is murkier than common discourse suggests.

Hypothesis 2: The Language of Approximation - A Toolbox for Psycholinguistic Communication

a. The language used to communicate measurements is chosen according to the intended receiver.

Playwright Jennifer Blackmer laments that, when judgements of value are made about a work of art, they are often “imprecise” and frustrating to the creator. Qualitative evaluations such as “That was the worst play I’ve ever seen” or “That movie was awesome” are easily tossed around, especially by those who are “not well-versed in thinking about art.” Refreshingly, Blackmer does not stop her exploration of public valuations there, and instead turns a typical grumble about careless critiques into an intriguing observation about the nature of language and its relationship to measurement. She notes that while “these imprecise evaluations...will never go away because they’re easy [for the public to make]....they’re the only vocabulary they have.”
My initial hypothesis on measurement and communication was focused primarily on what we can tell about the intended audience by the language the measurers chose for communicating their measurements.\footnote{To give an everyday example, I write in different language to myself if I know I will read it in a day as opposed to in a year, because I know my understanding tomorrow will be very similar to my understanding now. That cannot be said for my ability to understand myself a year from now: My note to year-from-now-me will be “clearer,” in fuller sentences, maybe even with a diagram, additions which would be unnecessarily bulky language for conveying the idea to tomorrow-me.} Contrastingly, Blackmer’s thoughts have more to do with the available vocabulary of the measurers themselves. In order to choose a particular vocabulary for an intended receiver, one would have to have access to that vocabulary in the first place! She points out that inaccuracies sometimes occur simply because the language necessary for useful evaluation are absent from the mental toolbox of the evaluator.

She then goes on to illustrate what is truly happening when a work of art is experienced: An attempted leap, across the gap between the artist and the audience, or between any two individuals who are “seeking to exchange meaning and truth with one another.” She concludes by suggesting that the true evaluation being made is not of the art itself; what we truly measure when we evaluate art is the communication gap, and how effectively that distance was crossed. This thought ends with a hopeful nod at her earlier frustration with the shortcomings of critical language: Evaluating the gap, if crossed successfully, “can lead us to a common vocabulary that doesn’t diminish the wants and needs of the artist, but provides him with valuable information to keep making art, which is something the world can never have too much of, right?”

To summarize: Blackmer’s essay disagrees with my hypothesis that the language of measurement is always chosen according to the intended receiver—because sometimes the people measuring don’t have those words in their repertoire to begin with. The language of evaluation is often not in terms which the receiver can find useful, but—as Blackmer affirms—it should be.

This line of thought circles back around to Thorson’s earlier assertion that description is a form of measurement, and that we “take the measure of the world by describing it.” I am tempted to repeat her thoughts on Kant’s denial of the correspondence theory of truth, and how that in itself represents another “gap” similar to Blackmer’s—a gap between the world as it exists and the world as we experience/perceive it. But, I must allow the reader to infer what connections can be made to previous sections, so that I do not continue to restate every essay’s arguments each time they reconnect to the topic at hand. I will point out, however, that Blackmer’s concern
about the absence of adequate vocabulary causing inaccurate/unhelpful evaluations could add another dimension to Thorson's questions concerning who gets to be a measurer.

b. Every measurement must eventually be interpreted by the human mind.

This hypothesis about measurement became important to me because of a punctuation mark: the comma. For example, consider the following article titles:

"0.5028193 x 10^7 Innocent Citizens Die in Terrorist Strike."
The absurdity of this representation is obvious, but perhaps the next will call attention to how subtle the linguistic representation of measurements can be:

"5028193 Indiana Citizens Voted in the 2012 Election."

Is that number low, or high? Our eyes search for the placeholders where commas can be inserted so that we can take in the number in terms of things we know, perhaps in terms of how many citizens live in our own city, or how many citizens live in Indianapolis. The commas, we assume, are present in numbers for accuracy, in order to reduce the likelihood of human error in copying them into databases, etc. In reality, as anyone who works with numbers in computer programs will tell you, commas are often a nuisance that must be eliminated through careful coding which will translate a number with commas to one without.

Punctuation in numbers, since the advent of advanced computing technology, is no longer used for accuracy in the calculable sense, but for accurate communication for human minds to estimate and take hold of. The commas are placed there for you—not your computer—for your comprehension and understanding, certainly, but also to influence your reaction to data. In this example, even the most "objective" form of measuring, the assignment of representative numbers, is primarily a form of communication intended to elicit emotional response and/or create cognitive connections which could not otherwise exist.

Historian James Connolly briefly touches on interpretation of measurements in his essay on history as a measuring of the human experience. The majority of his essay interprets "measurement" as "quantification," citing how historians have attempted to use empirical data (birth rates, voting totals, GDP, library checkout records) to piece together the untold stories of "workers, slaves, housewives, and others, the sorts of people who did not give speeches or write memoirs." In more recent scholarship, the use of "big data" is increasingly popular as a method of
gathering evidence about the lives of ordinary people. Yet, Connolly notes, this data is inevitably incomplete. Intermarriage rates, book check-out records, indeed—numbers, offer only skeletal understanding at best. I would add that, at worst, they can become the base for truth-claims which may go safely unquestioned because they are (seemingly) being backed by data. Do not mistake his argument—or mine—as undermining the value of quantitative research. On the contrary, Connolly describes his current work (with the What Middletown Read project) as primarily data-driven. To be sure, his essay only gave me more respect for quantitative historians, as sleuths of the scarcely heard stories of the marginalized in society, whose only voice in history lies hidden within the numbers. Yet, Connolly unflinchingly admits that "Knowing what someone read is not the same as knowing what they made of what they read."

This is where Connolly strikes a chord with my own observations: Numbers are seductively pure, promising to tell us objective truths about what we seek to know. And yet, eventually, any useful conclusions we draw from these numbers are extensions beyond the empirical data, into the realm of interpretation. Questions such as "Why did people do what they did?" and "What meanings did they assign to their behavior?" aren't answerable purely from a quantitative point of view, Connolly admits. Even further, he does not consider these questions to be wishy-washy or pointless for historians to pursue: "Answering these questions gets us to the essential, defining elements of the human experience. But they are subjective answers, the product of interpretation rather than objective measurement." This is an impressively honest assessment, particularly coming from a historian who has invested so much of his energy working to create a sizeable historical database. This shows a rare humility of perspective, acknowledging that quantitative methods are extremely useful for piecing together the past, but not self-sufficient for answering the most important questions about the human experience.

Connolly's essay agrees with my estimation of the value and shortcomings of empirical evidence and the danger of assuming that conclusions backed by quantitative data are in any way more "objective" or more reliable than conclusions reached through qualitative assessments. Every measurement, even those which seem to stand on their own, must eventually be interpreted by the human mind. Sometimes this is done by the researcher herself, forming an interpretation of the data which may seem to inarguably follow from her findings. Or, the data is simply presented to the reader, but in a way that suggests the interpretation the researcher wishes the reader to conclude. Even when the researcher is careful to present the information as absent of interpretation as possible, others can take the data and harness it for their own
purposes. Numbers can become a way of disguising subjective assessments, rather than a way of somehow moving beyond or above them.

This leads nicely into a discussion of a similar concern voiced in "On Choosing a Metric," by an Anonymous mathematician. He refers to "users" of mathematics, and their tendency to misuse metrics by not taking the time to understand their purpose, or by not considering the context of the data they wish to measure. (Indeed, in his essay, one could easily substitute every reference to "users" of mathematics with "abusers" of mathematics!) He praises "properly chosen metrics" and their ability to lead us to the important information we seek. Yet, he also notes that "misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what that metric really says, thinking that their heuristic ‘feeling’ of the data and metric is accurate and usable in some real-world interpretation they have in mind." He then goes on to give specific examples of when certain metrics do or do not make sense to employ in a given situation. Overall, his discussion seems to illustrate the two-part responsibility involved in measuring: The responsibility not only to measure accurately, but also to analyze the situation and understand how to choose a fitting metric to use in the first place.

Thus, the anonymous mathematician seems to agree with my concept of (even empirical) measurement as eventually subject to human interpretation. And even further, he views empirical measurements as being in danger of human interpretation. Yet, his argument is not so flat as to simply complain about human error and misuse of metrics. In his essay, humans are indeed the problem when it comes to accurate measurements, but they are also the solution. Humans, after all, are the ones with the power to assess the situation and educate themselves about what a metric can and cannot tell them, and choose their methods accordingly.

In his conclusion, he leaves with a somewhat surprising admonishment, one which seems to allude to more than just stock prices and jpg compression: "we should all keep in mind whether or not metrics were chosen for being most suitable to the aspects we wish to understand or whether they were chosen by default." Heeding this warning on a more abstract scale begs another glance at Thorson’s and Blackwell’s arguments about the default metrics that exist in social and academic discourses. The danger of metrics that are chosen by default, as the anonymous mathematician suggests, is that the resulting measurements may ignore context and lead to gross misinterpretations.

**Hypothesis 3: Measurement as the Act of Categorization**
a. All forms of measurement assign an object to a set.

This (and the following part, b) is perhaps the most important and most difficult to explain of my pre-project hypotheses, but I hope to clarify it as much as possible using examples that occur in the essays.

In his discussion of the role that measurement plays in the study of choice, economist John Horowitz discusses the poverty benchmark (or poverty line). Those whose incomes fall below this cut-off point are considered, by the federal standard of the word, to be in poverty. He then shows how complicated setting such a standard can be, particularly across cultures, family sizes, and sources of income. Although the issue of how to go about setting such a standard is a complex and worthy investigation, the questions I find most interesting about the poverty line as a form of measurement do not pertain to how the measurement should be formed, but how the poverty line itself works as a measuring device, separating the "poor" from the "non-poor." This effectively puts the population into two piles—piles that can cause a lot of controversy. After all, shouldn't we further distinguish between people who are poor because of illness, tragedy, or parental situation apart from those who are poor because of irresponsible behavior? Shouldn't we also distinguish between people who are working hard to better their situation, apart from those who take advantage of government "hand-outs?" Thus arises phrases such as "worthy poor," or the even more controversial, "working poor." The former distinction begs the question, "How does one determine who is 'worthy,'" and the latter raises a more systematic issue, "What causes economic conditions where 'working poor' no longer seems oxymoronic?" The distinctions we would wish to make in order to fine-tune what we mean when we talk about poverty only get stickier from there; the lines separating one pile from another get blurrier, and the battles over their exact boundaries become fiercer. Without these distinctions, contention arises against treating these piles as homogenous entities, when they unquestionably are not. Yet, the more distinctions we carefully attempt to add to an already problematic dividing line, the more convoluted it becomes to determine where the finer lines should lie, and which lines should not even exist to begin with.

On a more abstract level, it is useful to ask the question, "Is the poverty line a qualitative or quantitative measurement?" Certainly, it is a quantitative measurement, because it is determined by income. But, it itself certainly acts as a qualitative measurement, because it signifies that every member of the population whose income is below a certain point is
considered in "poverty." The measurements used to define the word "poverty" must be determined by qualities such as personal well-being and social inclusion, which can only be approximated by quantitative measurements. This is demonstrated by Horowitz's comparison of a person in the Congo living on less than $2 a day versus someone in the United States who needs about $32 a day. The costs required to be included in the basic social structure of a community vary wildly across cultures, even varying within the same country. After all, isn't a "poor" person in a city with excellent free public transportation and open community culture "better off" than if she were making slightly more money in a city whose culture was practically inaccessible to those without a car? Thus, quantitative measurements only make as much sense as they reflect the qualities that define comfort, inclusion, and capability of people to live well in a community.

When we set the boundaries of who can be considered to be in a certain "pile" or set, this action of measurement is contentious, and for good reason—it is extremely important to the people who are affected by these definitions, on both sides of the dividing line. (For further discussion, see Hypothesis 4(a), which explores how measurement affects the measured.) The poverty line highlights the correlation between qualitative assessment, quantitative approximation, and the resulting definitions that create the approximative category which we hope to make as accurate and useful as possible.

Sociologist Melinda Messineo discusses quantitative and qualitative measurements of society, believing that both have important roles to fill, each valid within the correct context. And often, the use of quantitative and qualitative methods may be simultaneously necessary for answering a single issue. She observes that, in order to study a group of people, you must talk to individuals within that group. This combines the methods that use "big data" (pertaining to the group) and qualitative assessment (collected from individual members and coded into categories). In Messineo's estimation, measurement is not only about assigning members to certain sets, but also about creating the set itself from information about particular members. In order to say something about (to use her example) minorities in advertisements, you would first have to ask questions which assign certain aspects of the advertisement into categories: "Is the character prominent? Attractive? Respected?" This part fits with my hypothesis that all measurement is the action of categorization. But Messineo offers an additional or alternate

---

3 I use "poor" and "poverty" in quotations not to imply that the federal standards are inaccurate, but to highlight the fact that when these terms are used in this context, they refer to a very specific, federally-defined measurement. They do not refer to our more intuitive interpretation of the words, though we might be tempted to unconsciously ascribe additional attributes to people whose incomes fall below that statistical marker.
hypothesis: Measurement is the act of creating a category out of individual bits of information, identifying some common quality among them. (For example, talking to individuals to see what commonalities can be generalized about the group, or noting a certain reoccurring pattern among commercials of a certain decade and giving a name to this trend.) Perhaps this can still be considered a form of categorical assignment, to a category which previously did not exist. Yet, still, this deductive reasoning towards a measure seems to extend beyond my former conception of measurement as categorization.

Blackwell (in his second response) makes mention of categorical measurement regarding the evaluation of student papers. Assessing student writing is both quantitative and qualitative, since it involves assigning a percentage according to how strongly certain qualities appear throughout the paper. Blackwell describes letter grades not just as symbols that happen to be attached to given percentages, but as categories of excellence (or lack thereof): “We can, with experience, support a claim that says most ‘C’ writing lacks certain features, or most ‘B’ writers demonstrate a similar level of mastery of rhetorical concepts like context or audience. With more experience, we can even refine such measurements according to pluses and minuses...” Thus, the numerical percentage assigned is actually just a refinement of where a paper belongs among the pile of other papers with similar qualities.

Blackwell claims that the evaluator’s accuracy in placing a paper into a letter-grade category increases over time, but offers no evidence for how this could be determined. How is it possible to know if one is becoming more accurate with grading papers over time, and not simply more comfortable and confident with the grades being assigned? This is not the only epistemologically problematic aspect of categorical evaluation. Blackwell notes that, although some attributes can seemingly be safely associated with a level of excellence, these attributes are not necessarily always present in every instance. There are always exceptions: “There will be at least one B+ paper (in the history of every B+ paper) that does not demonstrate a fair amount of in-depth critique...” I infer that, while some qualities can be reliably be predicted to appear in the B+ category of papers, there will always be an anomaly which has enough other excellent qualities that, though it lacks a quality typically present in a B+ paper, it still qualifies as a B+ paper.

This begs the question: Are these letter-grade categories descriptively defined or prescriptively defined? If they are descriptive, i.e., if one determines whether a paper belongs in a particular category by comparing it to the attributes of other papers previously encountered, it
would make sense that grader accuracy would increase over time. Time and experience would provide more papers that exemplify the attributes of each category, which would then make it easier to recognize where a newly encountered paper fits in with the rest. But if this is the case, then it would not make sense for there to be "exceptions" to the normal attributes of a B+ paper. If the categories are purely descriptive, it would be impossible to determine that a paper had a B+ level of excellence outside of this being evidenced by it having qualities similar to that of other B+ papers (or, its qualities falling short of A- papers and exceeding that of B papers). Exceptions to the norm imply that there is a higher, non-relative standard by which a paper is evaluated. How else could one tell if a paper is still of a certain level of quality although not meeting the relative standard, if not by appealing to a prescriptive standard? Yet, if that is the case, i.e., the categories are prescriptively defined, then there would be no need to evaluate papers according to how they compare to previous student papers, and grader accuracy (if we rely on the argument that accuracy is increased by recognition of certain features consistently appearing in a certain level of paper quality) would not increase over time. Prescriptive categories would make previous experience irrelevant.

The seeming dichotomy, between whether the qualities featured in student papers descriptively define the letter-grade categories over time or the letter-grade categories prescribe what qualities are required for a paper to make the cut, is actually a demonstration of how my original hypothesis and Messineo's additional observation work in tandem:

1. Measurement takes an object and assigns it membership in a set: This is a B+ paper. This evaluation can be achieved by how it fares in comparison with other papers, or by seeming to fit the qualities of a prescribed standard. (Or, through some combination of both relative and prescribed measures.)
2. Over time, measurement can also, in the process of (1), create a set of features which are common to the individual elements in the same category. This set can then be used to assess future candidates for (1): I've noticed that most B+ papers contain these particular features, and thus I will look for those features as a general guideline for determining whether future papers have achieved this status.

In this way, the qualities that often appear in each category work as a kind of check-list, a shortcut for proving to oneself that the paper does indeed belong in that particular category.¹

¹ Yet, this check-list is anything but absolute. For example, say that there are four qualities often displayed in B+ papers. It may be the case that each paper already in the B+ pile only displays three of the four (though it may be a different three for each paper). Thus, no absolute set of three (or four) qualities could be set such that every paper already evaluated as B+ meets the standard.
b. Assigning a scalable cardinality, which is what we most often we think of when we think of the word "measurement," is merely one example of the much broader action of defining sets and relationships between sets.

To put the above more simply, quantification is just one kind of measurement. Quantification is, like all ways of measuring, a way of signalling some kind of meaningful relationship defined by comparing, creating, or referencing sets. Numbers are just one of our favorite tools for representing these relationships. Even non-numerical metrics, such as American versus Non-American, could be represented by assigning a person a 1 or 0 value to represent their inclusion or exclusion as part of that particular set.

It is difficult to analyze how the essays respond to this portion of my hypothesis, since none of the essays discussed or attempted to define the abstract action of measurement. Yet, each essay acted upon some operative definition of measurement, and it is interesting to note which definition each assumed for its discussion. Below is a list of which essays agree with the basic premise that quantification is not the only type of measurement, which essays assume that measurement refers only to quantification, and which essays seem to alter their stance mid-essay.

**Broad Definition of Measurement**

John Emert (Mathematics)  
Jennifer Blackmer (Theatre)  
Patrick Collier (English - Literature)  
Juli Thorson (Philosophy)  
Todd McKinney (English - Creative Writing)  
Melinda Messineo (Sociology)  
Brent Blackwell (English - Literature, Holocaust Studies)

**Measurement Defined to Be Only Quantification**

John Horowitz (Economics)  
Anonymous (Mathematics)  
Mai Kuha (Linguistics)  
Paul Gestwicki (Computer Science)
Definition of Measurement Revised Mid-Essay

James Connolly (History)
- In the majority of the essay, "measurement" is used as if synonymous with empirical data. In the last paragraph, a broader interpretation of measurement is added. This switch is accompanied by the suggestion that (while exact measurement of the past is impossible) the broader we interpret what it means to measure human experience, the closer we get to a good approximation.

Jason Powell (Honors Humanities)
- The first half of the essay assumes measurement is quantification, but the last half of the essay develops the possibility that "measuring the trajectory of an idea" through qualitative evidence could be a form of measuring: "I have never thought of this process in terms of measurement, but I think it is. I measure the enlargement or diminishment of ideas."

Hypothesis 4: Measurement as a Two-way Relationship Between the Measuring Device and the Measurand, Each with the Potential to Affect the Other

a. The act of measuring alters the measurand.

This hypothesis has its origins in a puzzling quantum phenomenon: Measuring tiny particles changes them, sometimes changing their behavior, sometimes changing their location or velocity. This is intriguing because it means the measurement can no longer be assumed accurate, since whatever it measured has changed in the process of being measured. Likewise, I hypothesize that measurement often changes the thing measured (the measurand), not only when measuring objects on the quantum level, but in many other cases.

Creative writer Todd McKinney expresses anxiety about how measuring affects students—particularly young students like his son, who, upon entering third grade, embark upon the nine-year-long journey of standardized test-taking. McKinney's meditation on the subject of measurement is a continual stream, jumping from test-taking to woodworking to music to poetry without a breath in between. This medium seems to artistically reflect how life's anxieties and questions (no matter what sphere: work, home, or intrapersonal) tend to run together in the mind. It's interesting to note that, while connections can be made between the measurement problems he talks about, he himself makes no overt connection between them, leaving them for
the reader to infer (or not). He (indirectly) compares the problems with evaluating students through testing ("What if the student was having a bad day? Like her dog ran away?") with the problems he recently experienced attempting to measure a board on a table versus on the floor, and how the unevenness of the working surface can lead to inaccurate measurements. This begs the question, *What are we truly measuring when we measure students—their intelligence, the effectiveness of the test to reflect intelligence, or some function which depends on them both?* Mckinney also expresses concern about the fact that the scores his son earns now which will be attached to his ID number for the rest of his time in school: "How will teachers and administrators judge him?" I also wonder if we fully understand the extent of how constant testing/measuring of students affects the learning process and how students perceive themselves and their own intelligence—not to mention, how students are treated by peers, teachers, and administrators.

In his reflection on how faculty on the "academic stage" are evaluated, mathematician **John Emert** observes that "measuring always has an impact." He points out that, while many acknowledge qualities such as collaboration and engagement as important for university faculty to demonstrate, they are measured too often by "counting publications, presentations, patents, and grants." He argues that how we measure reflects what we value, and this can have a life-changing impact. While he shies away from specific illustrations of this "impact," perhaps in an effort to avoid a tone of complaint, I infer that he is alluding to how faculty can be negatively impacted by how they are measured. If some things (such as publications) are clearly valued, rewarded, and incentivized, those things will flourish in the university community. If other things, such as student engagement, community involvement, interdepartmental collaboration, are not celebrated and go largely unmeasured, those aspects may wane. In fact, faculty may even perceive themselves as effectively punished for engaging in these activities, since these take time away from other endeavors which are more effective towards building a prestigious academic career. Emert observes that “The most appropriate measurements may not be quantifiable or well-defined,” and even goes so far as to say that “perhaps it is not the measurement that is most important, but rather the act of measuring.” This last sentiment turns what could have been a purely negative outlook on how measurement affects faculty to a potentially positive one: Measurement is important—essential, even, because properly chosen measurements can motivate faculty to keep doing the things that make the university a place of learning, compassion, collaboration, and—as Emert points out—love.
Returning to Horowitz’s discussion of the poverty line, his essay concludes with two very bold questions concerning how people are affected by where the line is drawn: “Does the measure change people’s behavior? If money is given to reduce poverty and you can get more help if you are a single parent, will some people choose not to get married in order to retain benefits?” Economics, as the study of choice, is concerned with how certain policies will incentivize certain behaviors and/or discourage other choices. This is not unlike Emert’s concern with how (and for what) faculty are rewarded or, by lack of reward, effectively punished. Admittedly, the economic context is a much touchier subject, due to the fact that poverty discourse over the last 60 years has been riddled with harmful stereotypes that “victim-blame” the poor for being irresponsible or lazy, and thus presumably at-fault for their own circumstances.

Yet, abstracting this discussion from its highly politicized context, if we assume that people of any socioeconomic strata are indeed rational humans, all other things equal, it matters which behaviors are rewarded and which ones are not. To give an example, if a parent will receive $5,000 less in government transfers every year by accepting a work promotion only worth $4,000, which is the most responsible decision for the well-being of the family: to accept the promotion, or not? Clearly, the latter. These kinds of choices are not made out of laziness or irresponsibility—and possibly, just the opposite. This is not to suggest that welfare at large is not helpful and necessary, but to point out the importance of policies that are both compassionate and smart, providing healthy incentives. For example, the earned income tax credit (EITC) has support across the political spectrum as an effective policy for those measured as low- or moderate-income working citizens, incentivizing work while aiding struggling families and individuals.

Though Horowitz raises more questions than he provides answers, his concerns resonate with my hypothesis: We have to keep in mind that measurements do not just reflect society, they affect society. In fact, in my (and Emert’s) opinion, they are probably more sure to do the latter than the former.

Several essays spoke about the effects not of measurement, but of the lack thereof. In the following section, I address Collier’s mention of the changing literary canon, but I should note that Blackwell and Thorson raise similar concerns in their essays about the effects (particularly, the negative effects) of those who go unmeasured.
b. The measurand dictates (or should dictate) the measuring device/method.

When I attended the "Measuring Culture" conference at Notre Dame, I discovered that this was a foundational axiom (assumption) that all attendees held as true. That was an exciting moment for me, realizing that I was already 12 steps behind where they were with "my" hypothesis. They caught me up on the basics of their collective wisdom: If you conceive of culture as a discourse, you can measure it by talking with subjects. If you conceive of culture as something else, perhaps as an exchange of traditions, a collection of families, or as primarily a place and the activities that occur there, you will not be able to accurately measure it only from interviews. This reminds me of calculating partial derivatives of 3-D objects: Taking only one partial derivative will only give you a cross-section slice of the object—which can tell you something, but not much. The dimensionality of the object of interest tells you how it can be accurately measured.

Literary analyst Patrick Collier talks about how the standard literary canon has made a dramatic shift since the consciousness-raising movements of the 1960s. People had begun to question the New Criticism (and its emphasis on the "close reading of a relatively small number of canonical texts") that reigned during the mid-twentieth century. He notes that, "While this game-changing development began with a desire among students (and sympathetic professors) for a range of literary texts that more accurately reflected the real diversity of American life, it ultimately produced a rigorous and revolutionary theorization of literary value itself..." The dissatisfied students and professors who protested the lack of diversity in the curriculum ended up creating a philosophical debate among literary critics, who took the opportunity to reexamine the nature of literary value itself.

The question of "How do we determine/measure literary value?" is an interesting one, but, similar to my interest in the poverty line itself more than how it is determined or calculated, I am more interested in the canon itself as a measuring device. Why is it so important to students and professors that the literary canon is diverse? According to Collier, they sought a canon that "more accurately reflected the real diversity of American life..." suggesting that the literary canon acts as an approximation of a culture's lived experience, what we value, and what we wish to pass down to future generations as our cultural narrative. If the narrative leaves out the voices of women, minorities, and non-western influences, how might this affect students' perception of...
their cultural past, and as a result, affect our cultural future? The canon, as an approximation of American culture, must be dictated by what it intends to measure: the full, complex, diverse lived experience of all Americans.

This concern harmonizes with philosopher Thorson’s concern about who the measurers of society are. And, although the anonymous mathematician might also agree with the above concerns, he would probably criticize the wording of my hypothesis: In his essay, it is not the measurand which dictates the measuring device, but the measurer. After all, though an object of study can passively “suggest” how it can be best be measured by its own nature, it is the measurer who takes the time to understand the object and the possible metrics that can be used to measure it.

**Hypothesis 5: Measurement as Approximation**
Measurement helps us approximate what we cannot know exactly; it’s how we break-into-pieces/digest/understand the continuous/infinite/unknowable.

Thorson’s discussion of Kant’s denial of the correspondence theory of truth sheds a bit of light on this hypothesis. She would attest that “the unknowable” would actually apply to everything in the world, since the world we seem to experience is only estimated by our perceptions. Thus, through her lens, all measurement is by nature an approximation.

Connolly reiterates this in his essay on how historians have attempted to capture the past: “They believed that by studying birth rates, voting totals, gross domestic product, and other empirical measures, they could make sense of the collective experiences of workers, slaves, housewives, and others, the sorts of people who did not give speeches or write memoirs.” Though he freely admits that “human experience…is immeasurable,” he celebrates historians’ attempts to accurately recover approximations of it to the best of their ability.

Messineo echoes a similar aspect of sociological measuring: “in order to study a group, you have to talk to individuals,” although she also notes, “This dualism between the individual and the group is a challenge to the field.” Sociologists must rely on individual testimonies in order to get information about the group, a process that is, in her words, “potentially controversial in the field.” They understand the possible pitfalls of studying the whole in terms of
its parts, and as a result have “many mechanisms that we have to put in place to make sure we do not introduce bias into the process at any stage.”

Humanities scholar Jason Powell, unlike Messineo, is not in the least concerned with the possible problems introduced during the collection of qualitative data. In fact, just the opposite: He begins his essay with a humorous denouncement of quantification as a way of measuring human experience, quoting his Intellectual History graduate advisor: “we don’t fucking count anything.” After all, “How does one measure ideas?” he asks, noting the ways that numbers fall short of even approaching the task. He then points to the qualitative clues he looks for when tracking the progression of existential thought (his particular area of scholarship): “Sometimes it comes down to an author’s innocent statement in a letter or another piece of literature that lets me know that he or she arrived at a certain conclusion...In a sense, I try and measure the trajectory of an idea, see how it grows and flourishes or wanes and dies.” These bits of evidence work as qualitative scatter-plot data which shows the broader, intangible movement of conceptual development. Thus, the “murky” nature of measuring the continuous development of ideas becomes clearer, or at least more tractable, through being marked by the discrete collection of qualitative evidence.

Hypothesis 6: Measurement as an Implication
Every measurement is an if-then statement:
“If we are comfortable with this tool’s ability to assign our object of interest to a set (category) with an acceptable margin of error, then _____ is the approximative set to which the object belongs.”

This hypothesis could be stated more intuitively as “Measurement is a relationship between the object being measured and the proper calibration of the measuring device you are using.” My intention is to point out that measurement always requires starting assumptions, and the validity of the result will depend on the amount of error you approved beforehand. The reason I describe this as an implication (if-then statement) is to highlight a logical error commonly made during the measuring process.

In abstract logic, there are three parts to an implication such as \( p \Rightarrow q \), which would read “if p, then q” or “p implies q.” There is the p, which is the “if” statement, the q, which is the “then” statement, and there is the implication that forms the relationship between them. Assuming that the implication (the full statement) is valid, p being true would cause q to also be true. However,
if p is false, q could still be true and the implication would still hold, because the implication only forces q to be true if p is true. For example, “If I get my thesis done today, then I will bake you a dozen cookies,” is still a true statement if I do not get my thesis done and yet still bake you a dozen cookies. In fact, if the “if” condition is false, then anything could happen and the implication (full statement) would still hold as not being a lie.

In my application of this logic to the measuring process, p is the starting assumptions (or contextual calibration) of the measuring process, the implication is the measuring device, and q is the resulting measurement. I argue that resulting measurements have no bearing whatsoever if the starting assumptions are not verified. Some measurements are assumed to be “objective,” “empirical,” or somehow inherently meaningful without a given context. I assert that there is no such thing as a meaningful measurement without context, that is, without starting assumptions which are held as true. Of course, you could always have “objective” measurements which are always true no matter what your assumptions are. For example, I could tell you that this thesis paper is a 10; no matter who reads it, no matter what the weather is like, no matter how many years go by, it will always be a 10. But what does that tell you about my thesis, without any context for the scale I’ve chosen, or what the scale intends to measure about the thesis, or who has evaluated my thesis as a 10? In terms of providing meaningful information, this measurement is absolutely useless. Similarly, an implication’s latter “then” statement can still be true whether or not the starting assumption is true. Yet, unless we know the front end of the implication, that is, unless we can confirm the validity of our starting assumptions when measuring, any measurement that results is meaningless.

Although this specific logical issue was not directly approached in any of the essays, a couple writers alluded to similar concepts. For example, Thorson noted that measurements are only as accurate as the people making them. The anonymous mathematician’s essay might help to further clarify my hypothesis: He points out that there are several metrics which can be chosen for any given problem, and one could use one of these metrics (without calculation error) but still come up with a bad answer. In this case, the metric operated correctly, but the user chose the wrong metric to use in the first place. Either the context of the problem was ignored, or the user did not take the time to understand what the metric was built to measure, and thus the starting

---

5 In this example, p = “I get my thesis done today,” and q = “I will bake you a dozen cookies.”

6 It is only fails to be a valid implication if the “if” part is true, but the “then” part fails to be true. Thus, if I get my thesis done today but do not follow through on baking you cookies, then the implication is invalid.
assumptions were false. This is not unlike holding a compass north-side down, and getting lost as a result. The device isn’t the problem—the one misusing it is.

**Hypothesis 7: Humankind as Measurers**

a. Humans are the most accurate measuring device that could ever be devised.

This hypothesis was inspired by the words of German poet/scientist Goethe, who questioned Newton’s reductionist study of light (which divided light into its fractional parts):

> It is a calamity that the use of experiment has severed nature from man, so that he is content to understand nature merely through what artificial instruments reveal and by so doing even restricts her achievements...Microscopes and telescopes, in actual fact, confuse man’s innate clarity of mind.\(^7\)

Rather than place a device between himself and nature, Goethe elevated the powers of bare human senses and their ability to perceive: “The human being himself, to the extent that he makes sound use of his senses, is the most exact physical apparatus that can exist.”\(^8\) While it seems impossible that humans could be well-described as being better measuring devices than machines, consider that every machine created to measure must be verified for its accuracy by a human. A human may not be able to measure tenths of ounces as well as a scale, but a scale must be told when to recalibrate itself. Even a self-calibrating device must be programmed to know when it should do so. A human, like a ruler used to measure the temperature of water, may be famously errant (hence the phrase “human error”) when it comes to measuring certain things, but she is the best measuring tool for what she is meant to measure: measurements themselves. Thus, the creation of devices to measure does not remove the work of measuring from man, but in fact elevates the human responsibility to evaluate the effectiveness of methods of measuring.

> To this hypothesis, the anonymous mathematician would probably reply “No—and yes.”

As previously mentioned, he illustrates how the wrong “user” can be detrimental to the measuring process, by assigning an incorrect meaning to a measurement. And yet, his essay also argues that measurements should be chosen according to human perception, at times—for example, choosing an image memory metric according to how the human eye perceives an object such as a red square: “...when looking at a red square on a white background, the eye

---

\(^7\) Lehrs, *Man or Matter*, p. 111, p. 106.

\(^8\) Goethe: *Scientific Studies*, p. 311.
spends little or no time focusing on the large white regions outside the square nor on the large red region inside the square...The useful information in the picture is in the boundaries, and so the metric should take this into account."

This is similar to what linguist Mai Kuha observes about measuring the length of conversational pauses: "Pause duration can be measured with precision from a computer display; however, in many studies, especially in the early days of conversation analysis, stopwatch timing or even perceptual judgements were used. These low-tech methods may seem shockingly imprecise; and yet, depending on the researcher's goals, they have some benefits as well. In interaction, the objectively measurable duration of a speaker's pause may be less important than the listener's subjective impression of it..." Kuha illuminates the possible benefits to using human perception (as imprecise as it can be) to measure a phenomenon such as conversational pauses, for reasons not far from our previous discussion of how the method of measuring must be chosen according to the thing we hope to measure: "Pauses can influence how fluent a second language speaker seems to others." A conversational pause does not only consist of how long the person speaking was silent; it also consists of how long the other participant in the conversation perceives the pause to be. In order to measure something which will only fully exist once it has been interpreted through human perception, it makes sense to employ human perception in the measuring process.

b. The act of measurement is central to our humanity.

Reflect with me, for a moment, on the absurd amount of measuring and measurement a person undergoes within his or her lifetime. The first thing the doctor says, instead of welcoming you into the world, places you into a set: girl or boy. Minutes later, you are weighed carefully for your ounces, and even further, for your tenths-of-ounces. When you die someday, they will record the moment to the minute. They will carve the month and year into a stone, they will recount your life in brief oral histories, they will measure you head-to-toe in order to build a casket, slightly taller than you are. In between all of that, you will measure and be measured for all sorts of things—your weight, your money, your height, your intelligence, even your love. While the human obsession with measurement can have its excesses, our traditions, commemorations, as well as our explorative and artistic endeavors have a way of approximating the truth, the
divine, and the beautiful in ways that allow us to connect closer to one another and better grasp our role as humans in an universe of flux.

Though we might even lament the ways in which measuring can be abused or misused, we can’t escape the compulsion to evaluate and approximate in an attempt to understand human existence and experience. As linguist Kuha notes, even when simply speaking to one another, we are constantly measuring the absence of sound between words and sentences and using the results to deduce things about our conversational partners. Thorson’s verdict about the correspondence theory of truth accurately reflects our relationship to measurement as a whole: “None of us can get through the day without it.” We even measure things that we know and freely admit are subject to experience—such as art, which, as playwright Blackmer points out, is “impossible....but we can’t NOT measure art. We’re compelled to do it all the time.”

***

In the following section of my analysis, I outline some of the running themes, concerns, and observations regarding measurement that occur throughout the booklet, most of which were entirely new to me. Going forward from this thesis project, these themes offer new questions for me to explore as I widen my research to include scholarship beyond my own university.

I have left this portion of my project largely unprocessed, merely noting which scholars approached each theme and quoting the portion of their essay which touches on that aspect of measurement. I encourage readers to form their own interpretational bridges between these quotations, and evaluate for themselves whether the sum of the multiple perspectives on each theme is greater than wisdom of its individual parts.
Beyond My Own Hypotheses:
Measurement Themes that Arose from the Collection

1. Measurement Reflects the Measurers and Devices—Not Just the Measurand

- “We've ceased, in other words, to see greatness in literary texts as simply a property of the text itself and rather as a result of a complex and unpredictable set of variables. With our local power over syllabi and examinations, we are one, but just one, of these variables.” - Patrick Collier (Literature)
- “The discrete and formal measurements of Computer Science reflect its foundations in mathematics and electrical engineering.” - Paul Gestwicki (Computer Science)
- “What we see in the world and what we find worthy of measure is the result of epistemic and moral assumptions. These underlying assumptions structure how and what we measure so that, in general, those in power do well. Traditionally the tools with which we measure have been evidence, justification, and truth...all measurers are rational, unemotional, privileged, and most likely male.” - Juli Thorson (Philosophy)

2. Simultaneous Use of Multiple Measuring Methods; Multi-disciplinarity

- “Building any nontrivial software system requires a multidisciplinary collaboration, where the tools of the Computer Scientist are combined with approaches from the social sciences, humanities, and the arts.” - Paul Gestwicki (Computer Science)
- “Similarly, we often do research that has qualitative and quantitative elements. For example, I look at representations of minorities in advertisements...I then quantify those findings and run regression analyses on them based on various hypotheses about the media.” - Melinda Messineo (Sociology)
- “If disciplinarity is defined by the metric it uses to engage the world, then real world phenomena that exist in multiple states at the same time have to be reduced, diluted, and filtered. Interdisciplinarity Studies denies this singular outlook on studying the world. To truly understand and engage any phenomenon, it must be tackled on its own terms, which are never singular.” - Brent Blackwell (Literature, Holocaust Studies)
- “My interest in measurement developed over the last couple years, as I was toying with taking methods I was learning in one class and seeing if they helped me better understand problems in another class.” - Rebecca Jackson (Editor)

3. Purpose of Measurements
• "They find that formal processes and communication patterns, quantitative feedback, and technical structure give them the confidence to build and revise complex software systems." - Paul Gestwicki (Computer Science)

• "Holocaust Studies has a rarefied sense of precision with respect to numbers, dates and the like because of Professional Holocaust Deniers." - Brent Blackwell (Literature, Holocaust Studies)

• "It [measurement] promises to help us comprehend the past with precision and certainty." - James Connolly (History)

• "This is the purpose of measurement. To help us understand the world and make better choices. Because you can't always get what you want. Measurement helps us understand the tradeoffs that we must make." - John Horowitz (Economics)

4. Measurement Is Epistemologically Problematic

• "Kant claimed that the world as it is (the world in itself) is beyond our ability to know. All we can know is the 'phenomenal world' which is experience shaped by our organizing principles... We assume that we know the world because our descriptions have taken its measure. All of us act on measuring the world in this way. There are worries, however. What we see in the world and what we find worthy of measure is the result of epistemic and moral assumptions." - Juli Thorson (Philosophy)

• "[T]he biggest question of Computer Science—the question of whether P=NP—which is fascinating not only because we do not know the answer, but because we also do not know if it is answerable." - Paul Gestwicki (Computer Science)

• "The question of measurement is central to the field of sociology. We run on a continuum between the points 'the world is socially constructed and therefore accurate measurement is impossible' to 'the world is knowable and quantifiable and through this quantification predictions can be made.'" - Melinda Messineo (Sociology)

5. Is Objectivity Possible? Is Objectivity a Good Thing?

• "Measuring student writing is a notoriously tricky thing. For decades, composition instructors have struggled to assign an objective score to something that is essentially subjective." - Brent Blackwell (Literature, Holocaust Studies)

• "As a field, we [holocaust scholars] debate on the validity of the objective position itself, in as much as anyone can be in one. And to my knowledge, this is certainly one of the few fields where someone with an axe to grind or a personal grudge to bear, in fact, carries more weight than those of us who have none." - Brent Blackwell (Literature, Holocaust Studies)

• "Pause duration can be measured with precision from a computer display; however, in many studies, especially in the early days of conversation analysis, stopwatch timing or
even perceptual judgements were used. These low-tech methods may seem shockingly imprecise; and yet, depending on the researcher's goals, they have some benefits as well. In interaction, the objectively measurable duration of a speaker's pause may be less important than the listener's subjective impression of it..." - Mai Kuha (Linguistics)

• "Why was I so convinced that asking students to write papers and have conversations and craft well-formed ideas was a 'better' way to measure progress, especially considering that these assignments seemed harder for professors to objectively quantify?" - Rebecca Jackson (Editor)

• "Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind.” - Anonymous (Mathematics)

• "Is there an unbiased way to measure each other for how well we measure the world?" - Juli Thorson (Philosophy)

6. Misinterpretation of Measurements

• "Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind.” - Anonymous (Mathematics)

7. Immeasurability; False Precision

• “[O]ur ‘measurement’ of greatness of texts is anything but precise or absolute.” - Patrick Collier (Literature)

• "For the historian, measurement is seductive. It promises to help us comprehend the past with precision and certainty. Like all seductions, there is deception involved because human experience—the thing historians seek to capture and reproduce—is immeasurable.” - James Connolly (History)

• “Testing companies are even developing tests, I've heard on the radio, to record the speed with which a student answers a question, but also his or her hesitation. What if a student was having a bad day? Like her dog ran away?” - Todd McKinney (Creative Writing)

• “Martin Fowler—a recognized industry leader software development—has gone so far as to claim that productivity cannot be measured at all, and that any attempted measurements will only produce misinformation.” - Paul Gestwicki (Computer Science)

8. Exactness; Accuracy
• "Holocaust Studies has a rarefied sense of precision with respect to numbers, dates and the like because of Professional Holocaust Deniers." - Brent Blackwell (Literature, Holocaust Studies)

• "Simply put, accuracy is gained over time... The more papers someone grades, the more accurate the grading becomes." - Brent Blackwell (Literature, Holocaust Studies)

• "I have learned how to measure a poem, etc by reading a lot of poems, etc." - Todd McKinney (Creative Writing)

• "Because nature likes to repeat herself, we can get a much better approximation of the length of the whole by adding up tiny miniatures. This is similar to measuring the parameter of a fern’s silhouette by using a frond as your ruler." - Rebecca Jackson (Editor)

• "Pause duration can be measured with precision from a computer display; however, in many studies, especially in the early days of conversation analysis, stopwatch timing or even perceptual judgements were used. These low-tech methods may seem shockingly imprecise; and yet, depending on the researcher’s goals, they have some benefits as well. In interaction, the objectively measurable duration of a speaker’s pause may be less important than the listener’s subjective impression of it..." - Mai Kuha (Linguistics)

9. Bias

• "This process of taking a qualitative assessment and making it quantitative is potentially controversial in the field. There are many mechanisms that we have to put in place to make sure that we do not introduce bias into the process at any stage." - Melinda Messineo (Sociology)

• "Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind..."users" often don’t take the time and energy to learn what these metrics are (and are not) and can often make gross errors in their data analysis." - Anonymous (Mathematics)

• "Is there an unbiased way to measure each other for how well we measure the world?" - Juli Thorson (Philosophy)

10. Voice/Voicelessness

• "The material and cultural circumstance of those not in power provides a different way to describe the word, and hence, a different way to measure it. Not all measurements are acceptable. Some lack the evidence to claim that they describe something real in the world. The evidence of those on the down side of power, however, cannot be ignored if we
want to approach a more complete, and perhaps a more accurate, measurement.” - Juli Thorson (Philosophy)

• “They believed that by studying birth rates, voting totals, gross domestic product, and other empirical measures, they could make sense of the collective experiences of workers, slaves, housewives, and others, the sorts of people who did not give speeches or write memoirs.” - James Connolly (History)

• “Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind.” - Anonymous (Mathematics)

11. Data Doesn't Tell Us the “Why”

• “Counting votes or tabulating intermarriage rates leaves unanswered the biggest questions: Why did people do what they did? What meanings did they assign to their behavior?” - James Connolly (History)

• “Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind.” - Anonymous (Mathematics)

• “There are many quantitative historians, those who look at the numbers of whatever they are studying: crime, number of children per family, economic costs, etc. As an intellectual historian, in the words of my advisor in graduate school, ‘we don’t fucking count anything.’ ...How does one measure ideas?...[F]or me, measurement is not about quantity. It is a qualitative assessment as to how an idea works itself out. In a sense, I try to measure the trajectory of an idea, see how it grows and flourishes or wanes and dies.” - Jason Powell (Humanities)

12. Measurements/Evaluations that Respect Context

• “When we think of a book as great—and don’t be deceived, we do—we are apt to be thinking pretty specifically about what, for whom, and under what conditions others might perceive it to be great—and what exactly it is great for.” - Patrick Collier (Literature)

• “Should there be different standards [for the poverty line] in different countries? Do people in the Congo where poverty is defined as a person living on less than $2 a day need less resources than a single person in the United States who needs about $32 dollars a day?” - John Horowitz (Economics)

• “Pause duration can be measured with precision from a computer display; however, in many studies, especially in the early days of conversation analysis, stopwatch timing or
even perceptual judgements were used. These low-tech methods may seem shockingly imprecise; and yet, depending on the researcher’s goals, they have some benefits as well. In interaction, the objectively measurable duration of a speaker’s pause may be less important than the listener’s subjective impression of it..." - Mai Kuha (Linguistics)

Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind..."users” often don’t take the time and energy to learn what these metrics are (and are not) and can often make gross errors in their data analysis.” - Anonymous (Mathematics)

13. Subjective and Non-objective Measurements

"Where has this left us? Less certain of the validity of our own value judgements, certainly. Aware, as well, that whether a book (or a poem) speaks to a reader depends in part upon that reader’s experience....Aware that the literary value of any text is a process that occurs in history— a complex result of the interactions between writers, texts, readers, schools, and many evaluative voices that social media has only seemed to multiply in recent years." - Patrick Collier (Literature)

"Pause duration can be measured with precision from a computer display; however, in many studies, especially in the early days of conversation analysis, stopwatch timing or even perceptual judgements were used. These low-tech methods may seem shockingly imprecise; and yet, depending on the researcher’s goals, they have some benefits as well. In interaction, the objectively measurable duration of a speaker’s pause may be less important than the listener’s subjective impression of it..." - Mai Kuha (Linguistics)

While it may not seem as though these observations are actual measurements, per se, they are—qualitative measurements, not quantitative—and they evaluate the experience of one work of art set against others.” - Jennifer Blackmer (Theatre)

"The most appropriate measurements may not be quantifiable or well-defined." - John Emert (Mathematician)

"Kant claimed...all we can know is the ‘phenomenal world” which is experience shaped by our organizing principles.” - Juli Thorson (Philosophy)

"How does one measure development? The more I’ve lived, the less I know.” - Todd McKinney (Creative Writing)

14. How We Measure Reflects What We Value

"...this game changing development began with a desire among students [and sympathetic professors] for a range of literary texts what more accurately reflected the real diversity of American life...” - Patrick Collier (Literature)
• “What we see in the world and what we find worthy of measure is the result of epistemic and moral assumptions. These underlying assumptions structure how and what we measure so that, in general, those in power do well.” - Juli Thorson (Philosophy)

• “When students graduate, we would like to think they have become better thinkers, better citizens, more informed and capable. I’m not sure any educators hope students will simply be very excellent at memorizing many things for a short time and writing them on test sheets.” - Rebecca Jackson (Editor)

• “We reflect our values through our act of measuring: by what we measure, the means by which we measure, and the level of commitment which we dedicate to the process of measurement.” - John Emert (Mathematician)

15. Qualitative Vs. Quantitative

• “Pause duration can be measured with precision from a computer display; however, in many studies, especially in the early days of conversation analysis, stopwatch timing or even perceptual judgements were used. These low-tech methods may seem shockingly imprecise; and yet, depending on the researcher’s goals, they have some benefits as well. In interaction, the objectively measurable duration of a speaker’s pause may be less important than the listener’s subjective impression of it...” - Mai Kuha (Linguistics)

• “While it may not seem as though these observations are actual measurements, per se, they are—qualitative measurements, not quantitative—and they evaluate the experience of one work of art set against others.” - Jennifer Blackmer (Theatre)

• “[F]or me, measurement is not about quantity. It is a qualitative assessment as to how an idea works itself out. In a sense, I try to measure the trajectory of an idea, see how it grows and flourishes or wanes and dies.” - Jason Powell (Humanities)

• “[W]e have two broad ways of conceptualizing measurement: qualitative and quantitative. On some campuses, these conceptualizations divide the members into two camps and sometimes these camps are at odds with one another. There is often a sense that the quantitative work is more highly valued than the qualitative. However, on other campuses they are viewed as “tools in a tools chest” and whether or not a researcher uses a specific tool depends on the research question being asked.” - Melinda Messineo (Sociology)

• “When interviewed about what most contributed to their successes, teams inevitably split their answers between the measurable and the non-measurable. They find that formal processes and communication patterns, quantitative feedback, and technical structure give them the confidence to build and revise complex software systems. At the same time, they recognize that empathy has the most dramatic effect on both team productivity and the quality of the product. These two sides, the scientific and the humanistic, are not opposed to each other: they form a synergy that can propel a team to success or, in their absence, drag a team to a near standstill.” - Paul Gestwicki (Computer Science)
16. Who Gets to Be a Measurer?

- "[T]raditionally those who have measured assumed a uniformity; all measurers are rational, unemotional, privileged, and most likely male." - Juli Thorson (Philosophy)
- "[T]raditionally those who have measured assumed a uniformity; all measurers are rational, unemotional, privileged, and most likely male." - Juli Thorson (Philosophy)

- "In Holocaust Studies...there is a stark divide between those who experienced the event and those who study it. Ellie Wiesel is the most famous proponent of this division, believing that the ones who did not witness or experience the holocaust as such, have no right to an opinion about it...And to my knowledge, this is certainly one of the few fields where someone with an axe to grind...carries more weight than those of us who have none." - Brent Blackwell (Literature, Holocaust Studies)

- "Commonly, misunderstandings of statistics come from those who are willing to attribute a meaning to a measurement beyond what the metric actually says, thinking their heuristic “feeling” of the data and metric is accurate and usable in some real-world interpretation they have in mind...“users” often don’t take the time and energy to learn what these metrics are (and are not) and can often make gross errors in their data analysis." - Anonymous (Mathematics)

- "[T]here are people called critics whose primary responsibility is to evaluate art, and tell us whether or not a piece of art is worthy of our time. Which is also frustrating because the critic is simply one human being who responds to a piece of art on one particular occasion, and if she happens to be in a bad mood, or recovering from a nasty case of food poisoning, then her response will most certainly be filtered through her experience, whether she wants to admit it or not." - Jennifer Blackmer (Theatre)

17. What Happens When Measurements Disagree?

- "Because two [holocaust] survivors’ accounts differ with respect to what most would call objective or measurable data (dates, locations, numerical figures, etc.), they then declare that all testimony is suspect." - Brent Blackwell (Literature, Holocaust Studies)
Navigating the Next Steps

"Of all things the measure is man—of the things that are: that they are, and of things that are not: that they are not." - Protagoras

This thesis project has now reached its conclusion, but the questions about measurement that first inspired it have only multiplied. Creating *This Is Not a Pipe: Essays on Man as the Measurer of All Things* was an amazing opportunity, giving me an excuse to talk with scholars in departments on campus that I’d never even stepped foot in before. And, in the process of writing my analysis of how the faculty responses related to my own thoughts on measurement and how their essays “spoke” to each other, I got to play conductor to a host of skilled instrumentalists. This thesis was a great excuse to pull together some of the best minds I knew into the same “room,” so I could hear the harmonies and dissonance that result from different instruments asked to play along to the same song. After all, measurement is the music that all fields of inquiry play, albeit in different modes. I’m grateful that so many individuals loaned me their wisdom to work with, because the nature of measurement was something I couldn’t explore on my own—and maybe none of us can.

From my perspective, that is what we can take from Protagoras’s famous pronouncement of humans as the “measure of all things.” His statement is often interpreted as a radical relativism, a denouncement of the existence of objective truth. Eliminating the possibility of objective truth may offer a solution to this statement, but Protagoras presented to us something much less resolved and much more intriguing: a measurement problem. Rather than inferring the nonexistence of a reality beyond human perception, I let Protagoras’s statement stand as it is: measurement is subject to human perspective.

Any attempt to completely eliminate subjectivity or error in the process of measuring is to deny the role humans play in accurate (as possible) approximation: proper calibration of the metric according to a particular context (which includes the measurand itself), choosing the right starting assumptions, and carefully interpreting the results according to what the metric can really say. Often, the measurand cannot be well-approximated from one perspective, because a single metric cannot take into account all of the dimensions that deserve recognition. This is not to be lamented, but celebrated: measuring multidimensional phenomena must be a collaboration.

---

9 http://www.iep.utm.edu/protagor/#SH3b. (DK80b1). [Punctuation mine, for clarity.]
For every question that inspired this project, the process of answering them has given me the vocabulary to ask ten more. The last section of new themes that arose from the booklet was an unexpected addition, giving me a more comprehensive look at how this topic extends beyond what I originally estimated to be its breadth. The next phase of my pursuit will be to take the multidisciplinary conceptual vocabulary I've gained from this project and use it to investigate scholarship beyond that of my own university. Using my findings from this thesis as a starting framework, I can navigate previously unfamiliar fields, and have confidence enough in my basic understanding of measurement as a basis to build on.

This thesis does not conclude with a comprehensive definition on what it means to measure, and has not given me any hope that the pursuit of such a definition will meet "success," if we define success as completion. If, however, we define success as a continual process, a direction one travels toward what is worthy of pursuit, then I fully intend to keep putting one foot in front of the other towards the answers that will always be beyond me. Our duty as measurers is to approximate the best we can, to get closer and closer to approaching the impossible, the indefinable, the infinite, the beautiful—not because we can ever fully reach it, but because of what we learn about ourselves and our world when we try.

The thing I intend to study for the rest of my life is beyond and above me, something I'm sure I'll never fully comprehend—which doesn't worry me in the least. I know I'm in good company.
APPENDIX:

Project Invitation

The following document was given to each scholar who contributed an essay to This Is Not a Pipe: Essays on Man as the Measurer of All Things, providing them with four prompts on the subject of measurement. Participants were invited to write an essay in response to any one (or more) of these prompts—or to none of them, if they so wished.
"The moment you start to arrange the world in words, you alter its nature. The words themselves begin to suggest patterns and connections that seemed at the time to be absent from the events the words describe." David Shields, *Reality Hunger*

For my senior honors thesis, I am gathering a collection of miniature essays (each consisting of a minimum of one word and preferably no more than a few pages), written by scholars of various fields including mathematics, English, philosophy, computer science, history, with each approaching the topic of measurement. **This is your invitation to participate in this project, as an author of an essay on measurement.** Because this is a wide and open-ended topic which can be interpreted in an infinite number of ways, below is a list of four questions which may help to direct your response, should you elect to participate. At the conclusion of this project, participants will be given a booklet of the completed collection, and their names will be listed with their response unless anonymity is preferred.

It should be noted that this project is informal and creative in nature. The research portion of the project will be carried out by myself during the course of the semester as I write my own reflection on measurement. It is not necessary for participants to cite sources for information included in their response, although citations are always welcome if the author wishes to provide them.

You may respond to any one of the following questions. (Of course, you may include answers to more than one, but this is entirely up to you.)

1. What is the role of measurement in your field?

2. What does it mean to measure with accuracy?

3. Describe a metric within your field, or perhaps the standard "ruler" you use.

4. Describe a problem with taking good measurements, or a problematic effect of taking measurements.