
ACTA
COGITATA

An Undergraduate Journal in Philosophy

Issue 4

ACTA COGITATA

Issue 4
October 2016

EDITOR IN CHIEF
Professor W. John Koolage

STUDENT EDITOR
Katie Coulter

Submission Information: <http://www.emich.edu/historyphilosophy/journals.php>
Journal Home: <http://www.emich.edu/historyphilosophy/actacogitata/actacogitatamain.php>
ISSN 2330-5118

EDITOR'S NOTE: ISSUE 4

I frequently wonder what is the hallmark of philosophical thought. Whatever else it might be, a sustained and careful attempt to understand the world in all its coarse and fine grained detail seems a critical component. This year's authors provide us with no less than a robust philosophical take on issues ranging from the enduring damage of certain dualisms to concerns about the possibility of a grounding for moral judgements. I am, once again, struck by the sheer power of deep thinking and careful writing when it is brought to bear on the questions that interest those who have found themselves in the early phases of their love of wisdom. I am pleased to provide their work for your reflection, and equally pleased to reflect on these questions that capture the imagination of the newest members of our long and storied discipline.

Acta Cogitata continues to adapt to its institutional circumstances. I am sad to say goodbye to my two-year collaborator and student editor, Katie Coulter. I wish her well on her new adventures. I am also extremely pleased to accept her paper for publication in this year's edition. Her thoughtful work has made the journal so much stronger. Once again, the journal finds itself in new circumstances. I have no doubt it will continue to grow in new and interesting ways, while offering work of the highest quality to our readers.

Thank you to our authors, their mentors and teachers, and the institutions that encourage philosophic work as a valuable and worthwhile part of our human place in the world. Our human story would be less rich without your dedication and work.

Dr. W. John Koolage

MISSION AND PURPOSE STATEMENT

Acta Cogitata is dedicated to providing a venue for undergraduate authors of original philosophical papers to have their work reviewed and, possibly, published. Publication acknowledges the work of outstanding undergraduate authors, rewards their efforts, and provides a home for some thought-provoking projects. In line with this purpose, *Acta Cogitata's* authors retain their copyright so that they may continue to develop these projects. The journal, however, does not publish work that has previously been published elsewhere.

The journal accepts philosophical papers from all areas of philosophy and seeks to promote philosophical discourse in any area where such discourse may be illuminating.

The journal is published annually, in October.

CONTENTS

- 5 The Purpose, Praxis, and Future of Academia: An Exploration of Fichtean Approaches to Education
Eri Svenson, Harper College
- 14 The Science of Communication: A Bayesian Account of Communication Strategy Selection
Danielle Clevenger, Eastern Michigan University
- 23 No Soft Doctrine: Royce on the Problem of Evil
Brandon Wright, Grand Valley State University
- 33 Minds and Bodies: Early Modern Social Justice
Clare Áine Keefer, Eastern Michigan University
- 41 Predicting the Truth: Overcoming Problems with Popper's Verisimilitude Through Model Selection Criteria
K. Raleigh Hanson, Washington State University
- 51 Solving Frege's Substitution Puzzle: Analyzing it in Light of Descriptivism and Direct Reference Theory
Ayesha Rehman, City College of New York
- 59 Searching for Ethics' Grounding: A Case for Moral Feeling and the Human Relationship to Nature
Katie Coulter, Eastern Michigan University

THE PURPOSE, PRAXIS, AND FUTURE OF ACADEMIA: FICHTEAN APPROACHES TO EDUCATION

Eri Svenson, Harper College

Abstract

Following the publication of Schulze's "Aenesidemus", which detailed a skeptical critique of transcendental idealism, philosopher Johann Gottlieb Fichte found himself undergoing an "intellectual revolution". Having agreed with many of Schulze's arguments, he concluded that to preserve the spirit of Kantianism he would have to establish a foundationally reworked conceptualization of it. Fichte emerged from this process with a framework grounded upon the innovative claim that we should regard that there is nothing for us beyond our own consciousness, which creates both ourselves, and the world that we experience. Furthermore, because we create our own experiences, it is possible for us to access direct knowledge about our experiences through our "productive imagination", and gain knowledge through experience. This approach piqued my interest because prevailing academic approaches to knowledge production are based upon materialist assumptions, Baconian procedure, and production-based outcomes, often at the expense of qualitative and experiential procedures. Wondering if Fichte's philosophies might be able to offer alternative, more balanced approaches for academia, in this paper I participate in an exploratory process examining Fichte's perspectives on pedagogy, scholarship, and education. Beginning with the question: if he wrote on the matter, what were Fichte's perspectives on pedagogy? I discuss his relational pedagogy and the challenges he experienced balancing his students' autonomy with his position as an instructor. Next, I ask: did Fichte address the purpose of scholarship and education in the broader, social sense? I suggest that his writings conceptualize scholarship as a public good necessary for the progressive development of humankind. Finally, I reconsider Fichte's place in contemporary academia, wondering: where do we go from here and can Fichte help us get there? Ultimately, I argue for the relevancy of Fichtean approaches in addressing the problems facing academia today.

**The Purpose, Praxis, and Future of Academia: An Exploration of
Fichtean Approaches to Education**

I. Imagination, Experience, and Creative Genius: Fichte's "Intellectual Revolution"

Following its publication in 1792, G. E. Schulze's "Aenesidemus", which details a skeptic's critique of Kant and Reinhold's transcendental idealism, was forwarded for review to Kantian philosopher Johann Gottlieb Fichte. Though initially promising that his analysis would be promptly returned, he ultimately spent months toiling over the text as he wrestled with the realization that he found himself in agreement with many of Schulze's critiques. Specifically, among other claims, "Aenesidemus" interrogated Kant's concept of "the thing in itself", which he used to describe objects as they exist outside the limits of our own consciousness and, consequentially, outside of what can be considered knowable. If we cannot develop knowledge of truths outside of our own consciousness, however, how can we presume to know the thing in itself exists at all? By the same principle, Schulze also rejected Kant and Reinhold's argument that we can deduce knowledge of objective realities from our mind's representations of them, as that not only presupposes the existence of particular objects, but also an unknowable causal link between the "thing in itself" and our experiences. In addition, he critiqued Reinhold's establishment of consciousness as the "highest principle" of metaphysics, calling into question if philosophy could ever establish such a thing in the first place. Combined, Schulze concluded that these problems pointed to irreconcilable contradictions in transcendental idealism, which had failed to protect itself from skeptical critiques.

While Fichte conceded to the soundness of many of Schulze's arguments, he remained committed to the continuation of transcendental idealist philosophy. Schulze's critiques, however, called for him to undergo an "intellectual revolution", from which he emerged dedicated to preserving the spirit of Kantianism, but having concluded that to succeed in doing so he would have to develop a foundationally reworked conceptualization of it. By the time his "Review of Aenesidemus" was complete, it contained an articulation of the framework that would allow him to engage in a thorough undertaking of this project. Specifically, Fichte's revolution of thought lead him to take the innovative philosophical leap of arguing, contrary to Kant's consistent presumption of the existence of the thing in itself, that we should in actuality regard that there is nothing for us beyond consciousness. Because our minds are the creators of both ourselves and the world we experience, to speak of anything outside of them contradicts the spirit of transcendental idealism. Through making this claim, Fichte deactivated the contradictions that Schulze had highlighted in his critique.

From this innovation, he made two more claims. First, that because we are producing experiences for ourselves, we can develop experimental techniques that will allow us to observe and curate firsthand knowledge of how the experiences are being produced. Through these procedures, we can develop knowledge about representations without having to default to unfalsifiable concepts like "the thing in itself". Second, while Fichte agrees that consciousness is a poor "highest principle", he nonetheless defends Reinhold's claim that we ought to develop such a principle. To defend against skeptical critiques, however, we should not construct a principle that demands upon claims of fact, but on an act that we can each experience for ourselves. In other words, rather than a claim of *fact* about consciousness, we should aim to establish a replicable "highest principle" that speaks to the *act* of creating consciousness. For Fichte, this is necessary because if philosophy is to become scientific, it must be more than thought about; it must be experientially *done*. It is free, experienced discovery, rather than discarnate arguments, that should be at the heart of the discipline.

The establishment of this framework is also the place where my own exploratory project began. From my first encounter with Fichtean philosophy, I found myself fascinated with his claim that we should consider the use of our productive imaginations to be a source of truth about the world, rather than as a distraction from objective analysis. He insists that it is possible for us to, and that we in fact should, experience the merits of philosophical conclusions for ourselves, granting an exceptional amount of agency to individuals in the academic process. Not only are we actively creating our own conscious experiences, but we should also be involved in discovering truths about them. While this approach to philosophical knowledge production is inclined strongly towards the scientific in its focus on experimentation and duplication, it nonetheless remains a procedure that puts qualitative experiences and contributions at its center. For Fichte, in this regard the “creative genius” is something that should be encouraged and regarded as valuable, rather than dismissed as impractical or out of touch with true knowledge.

While this claim is interesting on its own, it was particularly noteworthy to me because contemporary academia takes a markedly different approach to knowledge production. Rather than centering imagination, “creative genius”, and qualitative experience in its scientific processes, academic research is organized around the aforementioned materialist assumptions, Baconian procedure, and production-based outcomes. Claims of truth from lived experiences are often considered, at best, insufficient, and at worst, inadmissible because this form of data is normatively considered to be more biased than data obtained through classical Baconian procedures, such as experimentation within the natural sciences. In addition, students are taught to be receivers of information and are often not encouraged to be active participants in its discovery. Knowledge is something that is externally adsorbed and parroted, not experienced or created.

While I cede that there is value in these procedures, it is troubling when they produce an uncritical valorization of the natural and quantitative sciences at the expense of qualitative and participatory discovery. Contrary to popular mythos, these “realer” sciences are not immune from bias and have been culturally shaped along with qualitative forms of knowledge production. Furthermore, if we were to completely accept this materialism, there would be no space for human freedom, rendering us mere products of matter and denying recognition of our exploratory agency. As I continued reading Fichte’s work, I often wondered if his approach to knowledge production might be able to offer more balance to these methods, offering alternative approaches for the broader academic community that would synthesize opportunities for scientific methodology and qualitative experiences. This inquiry is what I have sought to explore through this paper, beginning with the question: if he wrote on the matter, what were Fichte’s perspectives on pedagogy?

II. Language, Autonomy, and Manipulation: Fichte’s Relational Pedagogy

In searching for potential sources on Fichtean pedagogy, the first article that I found was written by Sean Franzel and titled “‘Welches Gesetz ist der Mensch in seiner Wirksamkeit?’: Pedagogy and Media in Fichte’s Encounter with Mesmerism.” Franzel’s piece examines the parallels Fichte perceived among the relationships between mesmerists with their patients, and instructors with their students. While the article offered rich and complex analyses on a multitude of topics, much of it initially seemed to only be peripherally related to my specific

inquiry. Fortunately, I was able to recognize on closer reexamination that nestled among this content was a direct and salient explanation of Fichte's pedagogical approach. The authors write that "For Fichte, education is first and foremost to awaken a student's ability to think actively and freely" (7), a perspective he grounded in a fundamental concern for freedom and autonomy. The paper goes on to further explain that Fichte specifically believed that reading only lead to passive learning, and so instead he opted to utilize interactional and experimental lectures, engendering his students' independent thinking through relational pedagogical approaches.

Given the content of Fichte's metaphysical philosophies, learning this was unsurprising to me, but it nonetheless was still exciting to imagine the possibilities for learning and creativity that such a classroom, if successful, would engender. While it was often implicit throughout his groundbreaking works, his pedagogical philosophy makes it clear how central themes of agency are to him, recognizing the inextricable link between freedom and the capacity to engage in authentic intellectual discovery. This is something that is ignored in many normative forms of instruction, which can be highly dictatorial and inflexible. Consequently, students never truly experience or develop a sense of ownership over their learning, making it more difficult for them to value, internalize, and contribute to the knowledge they encounter. They become trained to accept the status quo and struggle to become aware of the complexities of the world they experience and their place within it. By seeking to facilitate opportunities for guided, independent study, Fichte flips this script, respecting his students' creative processes, giving them the necessary tools to experience truth for themselves, and opening up the space for the dialectical process to continue.

Franzel's article does, however, describe instances in which this underlying philosophy was challenged and created inner conflicts for Fichte. After observing a mesmerist patient begin speaking in a manner that seemed out of her control under the influence of her practitioner's words, he became troubled about the potential manipulative power of spoken language. Rethinking his own belief in the freeing capacities of oral instruction, he was reminded of students who were successful at an activity while he was facilitating it for them, but who quickly lost that capacity once out of the classroom. The power of the relational authority inherent in his position became a tension for him as his deep commitment to autonomy and clashed with the practical realities of his profession. He taught so that his students could learn to engage in a process of free discovery for themselves, but he was bound to doing so in a way that seemed to inherently reduce students' autonomy.

While considering Fichte's struggles on this matter, I could not help but recognize the relevancy to the contemporary academic classroom. Regardless of the discipline, it has become a nearly universal experience to hear faculty express frustration over the difficulties many students have when expected to think and participate without direct and immediate guidance. Even with this guidance, if the activity asks students to take the reins on their rational capacities, they will often hesitate or even stop participating altogether. The possibility strikes me that many of these faculty likely share in the spirit of Fichte's aims in that they deeply respect and seek to bring out students' own critical thoughts but are then forced to reconcile that with students who struggle after being offered intellectual freedom. If an instructor accepts that at least one of the purposes of education is to foster students' freedom of thought, it seems as if it this would quickly become a pressing pedagogical challenge by creating a

conflict between one's intellectual, professional, and moral ideals, and the lived realities of the relationships between students, their instructors, and the institutions that form them.

Even based within my own experiences as a student, if I were asked to imagine what an ideal classroom experience would be, I can say with confidence that I would find Fichte's approach inspiring, but I am only cautiously hopeful about the possibility of it being realized on a broader scale after having watched a multitude of instructors experiment with similar approaches and come up against walls that were seemingly of students' own creation. If we sincerely are, as Fichte posits, agents of our own experience, why have we been so quick to deny it and cede to the manipulative power of others' language? Why do we shy away from embracing opportunities to take agency over our own experiences and knowledge? The roots of this phenomenon are clearly multifaceted, and I would not seek to deny the complex assortment of relational, intellectual, and sociocultural dynamics producing them. That being said, though Fichte's relational concerns did intuitively resonate with me, I could not help but think about what else might be shaping this experience.

What I kept coming back to is what, at least in contemporary times, we have been taught about the purpose of scholarship. Either as a cause itself or as a missed opportunity to generate solutions, the value of our inherent imaginative and creative capacities is rarely given experiential space. We are increasingly taught that our education is merely a means to more imminently "practical" ends, such as credentials and employable skills. The arts, literature, and humanities are demeaned, while students are pushed to enter technical fields regardless of their individual preferences or aptitudes. The purpose of the instructor and the classroom is no longer to assist students in fully experiencing their freedom, but to merely prepare them for more materially functional demands. The expectation that our education and work will be hierarchal and rote is normalized, and we dismiss as naïve those who aspire to more creative pursuits. These approaches clearly contradict Fichte's pedagogical philosophy and would prevent its incorporation throughout contemporary academia. Consequentially, I began to wonder: did Fichte provide a counter-perspective to this, addressing the purpose of scholarship and education in the broader, social sense?

III. Scholarship, Education, and Human Progress: Fichte's Scholarly Vocation

To begin exploring this question, I searched for sources on Fichte's "The Scholar's Vocation", a series of lectures articulating his perspectives on the social purpose of the scholar and their work. Written by David James, the article "Fichte on the Vocation of the Scholar and the (Mis)use of History" provides a discussion of these lectures. While James is specifically critical of Fichte's use of history as an instrumental tool, he offers a broader discussion of his philosophies, as well. According to the article, Fichte believed humans have been tracked onto a series of predetermined stages of development, which will ultimately culminate in a perfect reflection of rationality. At the time of his lectures, he believed that humans were on the precipice of entering a higher stage of the process, and that it was the scholar's moral obligation to grow a complete understanding of history and philosophy so that they could facilitate the elevation. For Fichte, the purpose of scholarship was to cultivate the capabilities of humankind as progressive beings.

This claim should not be mistaken to mean that he did not also recognize the pursuit of truth as its own end; as articulated in his pedagogical philosophies, such a value is in actuality at

the heart of his work. Consequentially, he would likely respond to James's concerns by arguing that his use of history as an instrument should not be taken as being a means to that end only or that its integrity would be sacrificed. Rather, his claim that the purpose of scholarship is in part to bring about human progress simply grounds the scholarly pursuit of truth in an additional, broader context. In its essence, "The Scholar's Vocation" articulates the social application of Fichte's metaphysical and pedagogical philosophies. Whereas in the classroom he sought to awaken individual students' critical capacities, here he positions that individual growth as one piece of a larger humanitarian puzzle. I consider the most meaningful contribution of this insight to be its recognition of intellectual exploration as a moral concern. The strengthening of our productive imaginations and the experiences they unlock are such an innate part of our humanity and such an inextricable part of our quest for human progress that there is an obligation to bring those skills and revelations to a larger, more accessible forum. Thinking back to his groundbreaking metaphysical philosophies, I would suggest this task should be regarded as a step in Fichte's project to reconceptualize philosophy as a participatory process, rather than a mere argumentative discipline, as it takes its conclusions into the public sphere where they can be experienced and applied. Through this, scholarship itself becomes a public good.

This provides a compelling counter perspective to the growing chorus of voices decrying the liberal arts as disconnected from the experiences and necessities of societies, claiming that liberal arts studies merely distract from "real" work. Of course, this is not to suggest that more "practical" disciplines, such as technology or trade work, are not of human value, as they are vital to the functioning and development of civilizations. Rather, what this is said to suggest is that the scholar's vocation should also be considered an uncompromisable pillar in the activities of human society. While building roads and innovating technology help us advance in production and trade, it is through the embrace of intellectual freedom and an imaginative pursuit of truth that we will progress in our humanity.

After "The Scholar's Vocation", Fichte gave a series of lectures titled "Addresses to the German Nation", through which he sought to inspire an increasingly demoralized and faltering nation. In the article "Fichte on Education", G. H. Turnbull highlights the central place that education occupied in this vision. While the lectures do contain a form of nationalism that in historical context we would now find troubling, I believe the core of his message can still be abstracted and breathed a more contemporary life. At their heart, Fichte's addresses advocated that education should be regarded an essential foundation for a successful nation, and consequentially should be made universally accessible to all regardless of their class. Rather than to create a skilled workforce, however, Fichte's educational vision sought to develop the innate capacities and characteristics necessary for humanity to reach progressively higher states. Relegating economic efficiency to a secondary role, the spirit of this system was a pedagogy based in our potential as whole, free, and imaginative beings, capable of taking ownership of those experiences if given the necessary training and opportunity. Thinking back to Fichte's trouble with his students, it is hard for me to imagine that a nation built upon such a foundation, valuing intellectual exploration as a necessary and universal right, would not be more likely to have developed the classroom experiences he was seeking for his students.

IV. Technology, Economics, and Changing Institutions: Fichte's Modern University

This brings us to my project's final question: Where do we go from here and can Fichte help us get there? While I would not posit to have a definitive answer, I can look back at where this exploratory process has taken me. Though I had always expected to find material that would contribute meaningfully to my understanding of the praxis and place of academia, I could not have anticipated it would be so imminently salient to the conversations we are having now. From his struggles in the classroom to his educational visions, it has become clear that the conflicts and potentials of Fichte's ideals are very much alive and poised for our present moment.

This is, furthermore, no coincidence. Writing in the *Chronicle for Higher Education*, Chad Wellmon reveals that Fichte was facing a landscape not entirely dissimilar to our own. At the time of Fichte's writing, the recent development of the printing press was making texts accessible to a degree not previously possible, rapidly expanding the percentage of the literate public that could access them. Prior to this point, universities had functioned primarily as "oral substitutes" for books that would have otherwise only been available in small, scattered libraries. With the printing press, however, this oral practice was no longer necessary, and consequentially the basic purpose of the university was being called into question. Many were arguing that academia should abandon its place as a home for ideas entirely, as they were no longer seen as economically relevant, and academia should instead transform into highly specialized vocational schools. Fichte, however, advocated against this, proposing a model that would place free, intellectual exploration at its heart: the university, according to him, should become a place where those with specialized disciplinary knowledge could go to teach, experiment, and create new ideas, rather than simply recite existing ones. This vision would become the basis for the university system we know today.

The internet has catalyzed society in previously unfathomable directions, making information and learning accessible in ways that it never was before. Idea exchanges that were once considered the hallmark of the academic institution are now happening throughout the web, and each new technological advance drives the need for workers skilled in developing and implementing them. As it was then, politicians, administrators, and consumers are increasingly calling us to step back from intellectual discoveries, marking the current model of instruction obsolete, and advocating for the creation of technical, skill-based universities. Even within my own institution, I have become increasingly conscious of and disheartened to hear conversations about innovative pedagogical proposals turn into nothing but an assessment of their impact on completion, retention, and employability. While these outcomes are undeniably important, the most meaningful academic experiences I have had, the ones that inspired me and caused me to grow as a thinker, have been those that have taken place in classrooms that sought to be laboratories for discussion and discovery, not those that imparted to me skills which I could list on a resume. Furthermore, as I have entered deeper into the workforce, it is the creative and synthetic abilities I developed in these classrooms that have most contributed to my professional success.

In his article, Wellmon argues that Fichte's innovations and their parallels to the contemporary moment should help us recognize the value in keeping academia as it is. By doing so, we will be able to ensure that scholarship's free exchange and creation of ideas is protected from the societal currents seeking to question its basic value. Though I concur with the importance of this latter mission, I respectfully disagree with Wellmon's conclusions. Rather

than as a reason to remain stagnant, the message of his article ought instead to be that, in the face of changing circumstances, we must remain willing and able to reimagine our purpose; if we don't, our home of ideas might be devalued entirely. With academia once more at this precipice of change, however, Fichte's philosophies should remind us that this reimagining does not have to come at the expense of our inalienable pursuit of creativity, imagination, and truth. Though we might be required to change, we should be thinking about how to do so in ways that allow us to preserve the spirit of our institution: institutions that value intellectual exploration for its own sake, while also recognizing that without intellectual exploration we cannot discover knowledge or imagine greater possibilities for human beings. Far from being distractions, in an information age increasingly overwhelmed with falsities and subsumed with bottom lines, preserving these values is more important than ever before. Rather than reimagining the university as a technical institute, we can reimagine it as a space that encourages opportunities for relational pedagogy, experiential pursuits of truth, and generations of new knowledge and ideas.

References

Franzel, Sean. "Welches Gesetz Ist Der Mensch In Seiner Wirksamkeit?": Pedagogy And Media In Fichte's Encounter With Mesmerism." *Germanic Review* 84.1 (2009): 3-25. *Academic Search Complete*. Web. 16 Dec. 2015.

James, David (2010). Fichte on the Vocation of the Scholar and the (Mis)use of History. *Review of Metaphysics* 63 (3):539-566.

Turnbull, G. H. (1923). Fichte on Education. *The Monist* 33 (2):184-201.

Wellmon, Chad. "In Defense Of Specialization." *Chronicle Of Higher Education* 61.30 (2015): B4-B5. *Academic Search Complete*. Web. 17 Dec. 2015.

THE SCIENCE OF COMMUNICATION: A BAYESIAN ACCOUNT OF COMMUNICATION STRATEGY SELECTION

Danielle Clevenger, Eastern Michigan University

Abstract

This paper will detail how Bayesian epistemology, traditionally a tool of philosophers of science, can be used to select a method of communication that is most likely to produce a desired communication goal from a targeted subject. Using the frame of Bayes' Theorem in the form of Posterior Probability Ratios, it will show how a communicator, focusing on agency and awareness, can use said frame to deliberately and purposefully select an evidentially favored communication strategy, intended to elicit a certain response from the respondent. By translating the epistemic version of Bayes' Theorem into a communication setting, this strategy presents an alternative method to use when navigating typical social interactions that would be useful for those who have trouble grasping traditional communication dynamics. Furthermore, the paper explains how this strategy is easy and natural to use because the human brain has evolved in such a way that it remembers and weights relevant occurrences for any given situation, which can then act as data for the comparative ratios.

The Science of Communication: A Bayesian Account of Communication Strategy Selection

Introduction

It has long been said that communication is an art, as opposed to a science. However, the psychological community has demonstrated that there are many underlying scientific principles that effect communication. This paper will detail how Bayesian epistemology, traditionally a tool of philosophers of science, can be used to select a method of communication that is most likely to produce a desired communication goal from a targeted subject. Using the frame of Bayes' Theorem, in the form of Posterior Probability Ratios, I will show how a communicator, focusing on agency and awareness, can use said frame to

deliberately and purposefully select an evidentially favored communication strategy, intended to elicit a certain response from the respondent. This strategy presents an alternative method to use when navigating typical social interactions. For instance, this method could help an individual decide if they should tell a friend that their significant other is cheating on them. This method encourages individuals to pause and review important evidence before deciding how to communicate in any given scenario. This pause and additional consideration enables individuals who are less adept at communicating to consider how their actions will affect the given situation, and it teaches them to make decisions based on how likely it is that the communication strategy will result in their goal.

Bayesian Epistemology

Reverend Thomas Bayes developed the mathematical theorem, which is derivable from a simple definition in probability theory, making it uncontroversial as a piece of mathematics (Sober, 2000). However, it has been applied in other areas, such as philosophy, where it becomes more controversial. The main premise of Bayesian epistemology is that knowledge and belief are not binary (i.e. true or false, known or unknown), but rather are a cohesive system of varying degrees of belief in numerous propositions. The degree of belief in any given proposition is known as a credence value, and credence values allow epistemologists a more nuanced way of considering traditional epistemic concerns.

Credence Values

It is commonsensical to understand that we hold some beliefs more strongly than others. For example consider two propositions:

A: The moon landing was faked.

B: My keys are on the table.

If asked which they were most sure of, most people would reply that they are most sure that their keys are on the table. Additionally, consider someone who has lost their keys. If asked where they believe their keys to be, they would probably say that they are the last place they can remember, such as on the table. The person may not be sure of this, but they say it because they believe their keys to be there *more* than they believe them to be somewhere else, like in the refrigerator (Koolage, 2013). Bayesian epistemology translates these degrees of belief into probabilities so that they can be used in various equations, just like a traditional mathematic probability.

These probabilities do not have to be the objective ones of traditional mathematics; in fact, most Bayesians take the probabilities to be subjective and often unique to the user. The job of the user is to then examine and employ these probabilities in their everyday life. Bayes' Theorem can be used as a tool that allows one to compare competing hypotheses to see which one is better supported by a given set of data. Bayesian Epistemology provides the user with a useful set of parameters that allows them to convert their belief hypotheses into an action guide that assesses the costs associated with acting on a particular hypothesis (Hartman & Sprenger, 2010). Bayes' Theorem also allows the user to introduce new data and see the effect

of this data on the probability of the hypothesis. In other words, users can determine whether the new observation makes the hypothesis more or less likely to be true.¹ The traditional mathematic version of Bayes' Theorem is as follows:

$$\Pr(A/B) = \frac{\Pr(A)*\Pr(B/A)}{\Pr(B)}$$

As an equation it states that, the probability of A given that B is true is equal to the probability of A multiplied by the probability of B given A is true divided by the probability of B. As previously noted, philosophers have expanded the use of Bayes' Theorem into epistemology because it allowed them a new manner of considering traditional epistemic concerns. This led philosophers to create an epistemic version of Bayes' Theorem that is as follows, where H stands for a hypothesis and O for a given observation:

$$\Pr(H/O) = \frac{\Pr(H)*\Pr(O/H)}{\Pr(O)}$$

Observation O is considered confirming evidence for hypothesis H if the $\Pr(H/O)$ is greater than the prior probability of H ($\Pr(H)$). The prior probability of H is the probability of the hypothesis prior to the consideration of the given observation. According to Bayesian epistemology, an often-updated belief has been through this equation multiple times, and the prior probability of the most current equation is the final probability from previous calculations. Using this equation, one can determine if an observation acts as a reason to believe in a hypothesis. More useful however, are comparative ratios, which allow the user to compare whether a given observation better supports one hypothesis over another. The Law of Likelihood states that a hypothesis is more likely to be true than a competing hypothesis if the observation in question was more likely to occur given that that particular hypothesis was true. Mathematically, that law is represented by the following inequality:

H1 is more likely given observation O iff $\Pr(O/H1) > \Pr(O/H2)$ (Sober, 2000).

Comparative ratios are derived by dividing the probability function for each hypothesis, and this results in the cancellation of the probability of the given observation. When all of this is done, the theorem looks like this:

$$\frac{\Pr(H1/O)}{\Pr(H2/O)} = \frac{\Pr(H1)*\Pr(O/H1)}{\Pr(H2)*\Pr(O/H2)}$$

Using this updated version of the equation, one can then decide, based on evidential support, which hypothesis is more likely to be true given the noted observation. In other words, is it more likely that Hypothesis 1 or Hypothesis 2 is true given that one has observed Observation O. Though one could conceivably determine objective mathematic probabilities for

¹ It is possible for evidence to neither confirm nor deny a hypothesis, but for simplicity's sake that will not be dealt with in this paper. For further information see Sober citation.

communication scenarios, this is impractical for daily life. Additionally, as real life cases are often complicated and convoluted, it is unlikely that each individual user would make the same determinations as another. For this reason combined with the fact that most Bayesians consider all credence values to be subjective, the probabilities are considered subjective. The step of using subjective credence values in the theorem allows one to consider using the theorem in the realm of communication.

The novel idea this paper details is the translation of the epistemic version of Bayes' Theorem, in the form of comparative ratios, into a communication setting where the communicator can use it to decide which course of action is more likely to lead to their communication goal. Seeing the hypotheses as communication strategies with varying consequences and the observation as the communication goal, one could then use data to decide which strategy was more likely to be observed if the desired response occurred. Furthermore, this paper explains how this strategy is easy and natural to use because the human brain has evolved in such a way that it remembers and weights relevant occurrences for any given situation, which can then act as data for the comparative ratios.

The next step is to now convert the equation into communication terms, as opposed to merely mathematic or scientific ones. In this case, one is deciding which communication strategy, if enacted, would be more likely to produce the communication goal based on prior evidence. The communication goal is now acting as a piece of data. The communication version of the equation would look like the following:

$$\frac{\Pr(\text{Communication Action 1/Communication Goal})}{\Pr(\text{Communication Action 2/Communication Goal})} = \frac{\text{Prior Probability of (C1)} * \Pr(\text{Communication Goal/C1})}{\text{Prior Probability of (C2)} * \Pr(\text{Communication Goal/C2})}$$

For the communication version of the equation, the hypothesis (H) will stand for the communication strategy being considered. The given observation (O) will be the communication goal, the desired outcome of a particular communication strategy, and the prior probability (Pr (H)) would be the subjective probability of the hypothesis. This probability would be based on all past data, taking into consideration various sources of data such as the success of the hypothesis in obtaining the response in the past, as well as other relevant information related to the given subject. This means that for communication purposes, the equation would read: The probability that the given communication strategy would result given the communication goal (Pr (H/ O)), is equal to the prior probability of the communication goal (Pr (H)) multiplied by the probability that the communication goal would occur given the communication strategy. This then allows the communicator to weigh two or more competing communication strategies to see which is more likely to produce the given communication goal. This would allow the communicator a logical, evidence based reason for choosing to communicate in a certain way. As opposed to the epistemic version of the theorem, which tells the user what to believe, the communication version tells the user how they should act, given the evidence and their communication goal. A simple example of the equation in communication terms would look something like the following scenario. Jenny and Sally are friends. Sally has discovered that Jenny's boyfriend is cheating on her. Sally doesn't know if she should tell Jenny or lie to her. Sally's consideration would then be this:

Goal: For Jenny not to be mad at Sally

C1: Lie to Jenny.

C2: Tell Jenny the truth.

$$\frac{\Pr(\text{Lie to Jenny/Jenny not mad})}{\Pr(\text{Don't lie to Jenny/Jenny not mad})} = \frac{\text{Prior Probability of (People being mad because of lies)} * (\text{Not mad/Lie})}{\text{Prior Probability of (Jenny being mad because of truth)} * \Pr(\text{Not mad/Truth})}$$

The communicator is determining if lying to Jenny makes it more probable that Jenny will not be mad at Sally, or if telling Jenny the truth is more likely to yield that particular communication goal.

The Prior Probability of the Hypothesis

The first part of the equation requires that the prior belief in the hypothesis be taken into account. For the purpose of communication, determining antecedent belief involves looking at how often the given communication strategy has produced the desired response in the past. Figuring this out into a quantifiable number is possible to some degree, but for practical communication purposes is unnecessary as the brain has evolved a system for ranking and weighting relevant occurrences. The brain naturally extrapolates and stores facts about how every person you have interacted with has dealt with certain situations. (Wilson, & Sperber, 2004).

There are two factors that must be weighted when determining the prior probability of a given communication strategy. The first is the number of times the strategy has produced the desired output in the general past. Meaning, how often the communicator is aware that sometime, somewhere, the given strategy has produced the desired behavior. The second is how often the strategy has produced the desired response in the given individual. The data that one has access to for a given individual is obviously increased and more accurate the more one interacts with that individual. That is to say, one will have a higher rate of accuracy in determining the probability that a given communication strategy will succeed with a well-known subject. Even with limited past scenarios, where there is restricted data either on the situation, subject, or both, the human brain is capable of using general past knowledge and small details to assume a reasonably accurate degree of prior belief. It is also interesting to note that most subjective Bayesians would argue that the prior belief doesn't have to be very accurate at all, especially in the initial application (Koolage, 2013). This makes the equation very practical to use in everyday life. Because human interaction is closely tied to factors such as character, values, and norms, even without having the advantage of previously encountering a certain situation one could make a reasonable prediction as to a degree of belief that the given communication strategy would produce the desired communication goal in a given individual. For an example of this, remember Sally and Jenny. Sally has found out that Jenny's boyfriend is cheating on her. Sally doesn't want Jenny to be mad at her. She can lie to Jenny and risk Jenny being mad that Sally didn't tell her, or she can tell Jenny and risk her being mad that Sally told her bad news.

In this instance the consideration looks like this:

Goal: For Jenny not to be mad at Sally.

- C1: Lie to Jenny.
- C2: Tell Jenny the truth.

Given that Sally has had a conversation with Jenny about how much Jenny disapproves of plagiarism, Sally is able to deduce that Jenny highly disapproves of people who aren't honest. Therefore, Sally can assign a high prior belief that telling Jenny the truth is less likely to make her mad than lying to her, even though the pair has never encountered a similar situation.

The Probability of the Observation Given the Hypothesis

The second portion of the equation is the probability of the observation given the hypothesis. Following the above example, this means that one would obtain the probability for C1 (lying), by determining how probable it is that Jenny would be mad at Sally if Sally lies to her. For C2 (honesty) one would determine how probable it is that Sally telling Jenny the truth, that her boyfriend is cheating, would make Jenny mad at Sally. This part of the equation would involve only general historical considerations. This is because, in order to maintain mental cohesion, the communicator must be able to separate how logical it is for a desired response to happen given the strategy, versus how probable it is that the response is going to happen given the individual.

Constructing A Strategy

Now that it is shown how one could use Bayes' Theorem with communication strategies, it is worth looking at how one arrives at a given hypothesis. It is important to note that for the purposes of this paper, one is using this method of communication as a high level meta-cognitive strategy for making decisions. This means that the individual using the method is considering all of the following information carefully and consciously. Constructing a strategy is then a very deliberate action. Constructing a viable communication strategy necessitates that the user consider the end goal(s), and what action seems to resonate with achieving these goals. The computing of the Bayesian Posterior Ratio then further pushes the user to consider the important past data they have regarding the possible actions. It is this pause to compute that is valuable in practical applications. It requires the user to stop and consider what data, if any, they are considering and if this data is relevant to the situation at hand.

Employing the Chosen Communication Strategy

To increase competence in any communication situation, the communicator must not only know what communication strategy to select, but also how to enact it, once it has been chosen. Knowing how to say or do something, with what phrasing, tone, timing, and place, is just as important in the process as knowing what to do. This means that once a communicator has chosen the favored strategy, they must then decide how to employ that strategy. Communication is a complicated give and take process that involves not only explicitly stated content but also implied content and non-verbal cues, such as body language. "The meaning of a word comes entirely from the word, and entirely from the speaker, although only the latter controls the context which helps determine the word's effective meaning..." (Hamer, 1970). It is this control of context that the communicator must be well aware of as they employ the chosen

communication strategy. An error in context determination can result in a well-chosen, evidentially supported strategy that does not result in the given communication goal. A highly simplified example of this importance can be seen in the following case. Alex has broken Kelsey's glass bowl that was left precariously on the counter. Alex wants Kelsey to forgive her for breaking the bowl.

Goal: Forgiveness

C1: Alex should apologize.

C2: Alex should say nothing.

Alex has determined that the probability that Kelsey will forgive her for breaking the bowl is much higher if Alex apologizes. However, this is highly dependent on the manner in which Alex apologizes.

Goal: Forgiveness

C1: Alex apologizes sincerely.

C2: Alex apologizes insincerely.

The probability that Kelsey will forgive Alex is just as dependent on how Alex apologizes as it is on the fact that she apologizes. This is a tricky, added layer that must be considered by the communicator when seeking and implementing a communication strategy that will produce a desired communication goal.

Implications

This method of considering communication certainly creates questions. Some of these questions are areas for further research and some of them are ethical questions that are also worth further inquiry. The claims made in this paper are intended to be normative rather than descriptive. Further research could be done in determining more clearly how a given individual weights competing stimuli; Sperber's research doesn't explain what mechanism does this, or how it does this. Ethically, the communicator using the equation would need to have high standards. There is no question that using Bayes' theorem in this manner represents an element of manipulation. A practiced and sensitive user would be capable of using his or her considerations to deceive or elicit a response that could cause harm. This means that if taught, there should be emphasis placed on choosing constructive as opposed to destructive goals. A truly constructive goal would benefit both the user and the targeted respondent.

Despite these concerns, there are numerous instances in which individuals could benefit from considering communication in this manner. First, it would be beneficial in general, everyday life. Many disputes are caused because there was an error in what was communicated. This method forces the user to carefully examine multiple facets of communication. The user must identify the end goal of their communication, question what is relevant, determine the probability that a given action will occur, and choose how to enact a chosen strategy. This greatly reduces the chances that an unintentional mistake will be made. Additionally, this strategy could be used to explain communication to those who have trouble grasping it in more traditional forms. An excellent example would be a subject that suffers from

autistic disorders such as Asperger's Syndrome. They could be taught to look for simple relevance points, and then considering their goal could construct strategies that they could then consider and possibly implement. There is a lot of good that can come from combining disciplines and considering communication in a new framework.

References

Andrews, Kristin. *Do Apes Read Minds?: Toward a New Folk Psychology*. Cambridge, MA: MIT, 2012. Print.

Hamer, Colin. "Meaning Things in Words." Ed. Dermot Moran. *Philosophical Studies* 19.0 (1970): 5-10. Web.

Hartman, Stephan, and Sprenger, Jan. "Bayesian Epistemology." (n.d.): 1-19. 17 Aug. 2010. Web. 5 July 2014.

Koolage, William J. "Likelihood Theorem." Eastern Michigan University. 28 Oct. 2014. Lecture.

Ramsey. "Truth and Probability." *Hugh Mellor Foundations: Essays in Philosophy, Logic, Mathematics, and Economics*. (1926): 58-100. Web.

Sober, Elliot. "Introduction to Bayesian Epistemology". 31 Jan. 2000. Teaching material from University of Wisconsin website.

Wilson, Deirdre, and Sperber, Dan. "Relevance Theory." *Handbook of Pragmatics* (2004): 249-280. Web. 24 Nov. 2014.

NO SOFT DOCTRINE: ROYCE ON THE PROBLEM OF EVIL

Brandon Wright, Grand Valley State University

Abstract

Because of the ubiquity of evil, religious systems, which aim to influence the way we live our lives, must answer three questions: what is evil, why does evil exist, and how can we eliminate, or at least manage, evil? Call this the *broad* problem of evil, as opposed to the traditional *narrow* problem of evil. I reconstruct the answer to the broad problem of evil found in Josiah Royce's later writings in the second section of this paper. Then, I explain why traditional theodicies are deficient answers to the narrow problem of evil. I argue that Royce's answer to the broad problem of evil merits a response from philosophers in the Abrahamic traditions because, while it is theistic—and even teleological—in nature, it does not presuppose the Abrahamic conception of God, nor does it suffer from the deficiencies of traditional theodicies.

No Soft Doctrine: Royce on the Problem of Evil

I. Introduction

The problem of evil presents one of the most serious challenges to the Abrahamic conception of God. The problem arises when two statements are conjoined: (1) if God exists, God is omniscient, omnipotent, and omnibenevolent, and (2) evil exists. If God is omniscient, then God must know if evil exists. If God is omnipotent, then God has the power and ability to eliminate that evil, given the desire to do so. And finally, if God is omnibenevolent, then God must desire to eliminate evil, or, at the very least, all unnecessary evils. Yet, evil exists. This seems to imply that God either does not have the three traditional attributes as defined or does not exist.² Call this the *narrow problem of evil*.

However, evil—henceforth denoting *undesirable states of affairs*—affects everyone, no matter their religious beliefs. Since evil is a major aspect of the human experience, religious systems must shoulder the theoretical burden of explaining it. Three major questions stand out:

1. Michael Tooley, "The Problem of Evil," in *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta, ed. accessed July 16, 2015, <http://plato.stanford.edu/archives/sum2015/entries/evil/>, Introduction.

what is evil, why does evil exist, and how we can eliminate (or at least manage) evil? All religious systems should provide answers to these questions, which may collectively be termed the *broad problem of evil*, whether or not they presuppose the Abrahamic conception of God.³ One system that answers these questions persuasively and does not presuppose the Abrahamic conception of God is the philosophy of religion proposed by Josiah Royce (1855-1916).⁴

In the next section, I reconstruct the answer to the broad problem of evil that Josiah Royce offers in his later writings.⁵ In the third section, I outline two traditional theodicies⁶ and argue that they are deficient responses to the narrow problem of evil, drawing inspiration from Royce's mid-career essay "The Problem of Job." I then argue that philosophers in the Abrahamic traditions should address Royce's answer to the broad problem of evil because it seriously challenges the status quo in Western philosophy of religion. First, it does not presuppose the traditional conception of God, while remaining theistic. Second, while it does not suffer from the deficiencies of some traditional theodicies which try to justify God's decision to create a world with evil, it still provides a teleological account of the existence and resolution of evil.

II. Royce's Answer to the Broad Problem of Evil

I mentioned that Royce does not presuppose the traditional conception of God and that this has major implications for his answer to the problem of evil.⁷ We will discuss Royce's conception of God toward the end of this section, after laying the foundation of Royce's views on evil. Royce understands evil in the typical sense but with a pragmatic twist. According to Royce, "evil" denotes any undesirable state of affairs which serves to undermine the purposes of a rational agent.⁸ We can also derive a definition for moral evil from this: *any undesirable state of affairs caused by an agent's willful action or inaction which serves to undermine the purposes of a rational agent*. Royce's commitment to pragmatism complements, rather than

³. Philosophers who attempt to answer the narrow problem of evil will find themselves answering the broad problem along the way, though they may simply take the answers to the "management" question for granted from the particular tradition they are working in. For example, a Christian philosopher who proposes a soul-making theodicy will certainly define evil and say why it exists (i.e., to facilitate the soul-making process). That same philosopher may then implicitly or explicitly defer to the Bible for its teachings on coping with evil.

⁴. I will not speculate as to whether or not Royce's conception of God can be considered a non-traditional, but still *Christian* conception of God, though Royce seems to indicate that he thinks this is the case.

⁵. For the reconstruction of Royce's answer, I limit my inquiry to *The Sources of Religious Insight* (1912) and *The Problem of Christianity* (1913).

⁶. See Tooley, Section 4. A *theodicy* is an attempt to give a reason why God (traditionally construed) might allow evil to exist, and how God could remain omnibenevolent in doing so. This is in contrast to a *total refutation*—an attempt to prove that the existence of evil does not offer even *prima facie* ground to argue for the non-existence of God. Also, a full explication of the distinction between a communal understanding of the problem of evil and an individualistic one would be much toso long for the present work. Instead, this discussion will focus on the first novel feature of Royce's answer: the non-traditional conception of God.

⁷. Jacquelyn Ann K. Kegley, *Josiah Royce in Focus*, (Bloomington, Ind.: Indiana University Press, 2008), 80.

⁸. Josiah Royce, *The Sources of Religious Insight*, (Washington D.C.: Catholic University of America Press, 2001), 216.

conflicts with, the standard definition of evil since God is a rational agent and acts of moral evil ostensibly undermine God's will in the world (or at least, God's intentions *for* the world). Royce does not defend the Abrahamic conception of God from the narrow problem of evil. Still, he agrees with traditional monotheists that people exist in a fallen state, fall short of a definite, ideal life, and need a savior to achieve that life.⁹ However, on Royce's view, people are not evil by nature. They perform evil actions because they are morally-detached individuals. In other words, each individual has interests, goals, and desires which can objectively conflict with those of others.¹⁰ If left unrecognized and unattended, moral detachment leads people to ignore the needs of others and to take so much pride in their own strivings that they fail to see the value in other's conflicting strivings.¹¹ Royce calls this unhappy state "social blindness."¹² In order to find the cure for the affliction of social blindness, we must investigate the origin of the morally-detached individual.

People are morally individuated in three ways: by the distinctness of their experience, the outward inaccessibility of their thoughts and intentions, and by the presumption that "deeds and their doers stand in one-one correspondence,"¹³ or, in other words, people presume that collective action is merely the sum of individuals' actions. Royce argues that this last idea is of recent vintage and is not supported by experience in daily life.¹⁴ On his view, a community is a superhuman being that is composed of, but not reducible to, its members. By extension, that being's actions are more than the sum of its members' actions.¹⁵ Members of a community overcome their moral-detachedness by uniting in the "spirit" of their community. That is, by taking up shared values and purposes and acting in the world together as one. These Roycean communities come in various sizes and persuasions, so we have plenty of candidates to choose from for an illustration. A hypothetical youth soccer league will do nicely.

Imagine that, some years ago, a group of parents decided that the local neighborhood children should have more opportunities to play together and get to know each other. They pitched in to buy a vacant field and soccer equipment and started holding games every weekend. Those parents formed a community by acting together for the sake of a shared purpose. Each member now considers the past actions of the league as events that belong to their own past, and the future actions of the league as part of their own future. For example, Bill and Sarah both remember painting the lines on the field before the very first game, and look forward to the day when the league can afford a scoreboard. Like an individual person, a community acts in the world presently, has a past, and will have a future. The sum of all those shared and anticipated experiences—and the meanings those events hold for the members—

⁹ Ibid., 28-29.

¹⁰ Josiah Royce, *The Problem of Christianity*, (Washington D.C.: Catholic University of America Press, 2001), 194.

¹¹ Kegley, *Josiah Royce in Focus*, 93.

¹² Royce, *Problem of Christianity*, 378.

¹³ Ibid., 238.

¹⁴ Ibid., 240.

¹⁵ Ibid., 123.

constitute the “self” of the community.¹⁶ Yet, not every community has the same peaceful existence as our hypothetical soccer league. Often, the purposes of a community are subverted by—or existentially opposed to—instances of evil, and members must thwart that evil. Royce calls people’s practical devotion to a higher communal life—including struggling together against evil— “loyalty,”¹⁷ and it is to that concept that we turn now.

Recall that an instance of evil is any undesirable state of affairs that undermines the purposes of a rational agent. Under this heading, we would do well to include pain, disease, and pestilence. Finite beings can only survive within a very narrow range of acceptable conditions and are severely limited in their abilities to maintain those conditions. Since human beings are limited in their individual experience and knowledge of the world, they form communities to survive. As a matter of course, those communities create and follow moral codes that vary widely. This means that the practices of one community are often considered evil by another. How can we reconcile competing and diametrically opposed loyalties that are equally moral on their own terms? And how can we do so without embracing relativism? Toward this end, we will need a regulative principle: a principle which is logically prior to the moral code of any particular community, but that every community can act upon. Furthermore, no community should have reasoned grounds to reject our sought-for principle, no matter how fiercely they might oppose any other particular community.

That principle is this: recognize “the spiritual unity of all the world of reasonable beings” as the true cause of loyalty.¹⁸ Then, seek to actualize that unity through the particular causes that make up one’s communal life. It is necessary and honorable to devote oneself to one’s community, but it is necessary to ensure that one’s community contributes to the overall growth of loyalty. Communities that are rooted in hating and destroying other communities are not objects of genuine loyalty. Those communities retard or reverse the overall growth of loyalty, because they exist solely to divide people from one another. They explicitly undermine the true cause of loyalty, which, as we will see, is divine in nature. On the contrary, so long as a person is loyal, and that loyalty is not given to a hateful cause, that person is doing their moral and spiritual duty.¹⁹ At this point, it would be tempting to infer that the highest moral life is a single-minded campaign to rid the world of evil. Yet, Royce’s keenest insight into the experience of evil is that this is blatantly false.

So far, we have been treating evil as something that simply should not exist. That seems intuitive. After all, curing 100% of malaria cases is necessarily better than curing 99% of malaria

¹⁶. See Chapter 2 of Kegley’s *Josiah Royce in Focus* for a thorough explication of Royce’s views on the self.

¹⁷. Royce, *Problem of Christianity*, 269. See also: Frank Oppenheim, *Royce’s Mature Philosophy of Religion*, (Notre Dame, Ind.: Notre Dame Press, 1987), 142. “Higher life” here is meant in two senses. First, in the sense that devotion enriches a person’s life, and, second, in the sense that the person’s life becomes more attuned to the divine life.

¹⁸. Royce, *Sources of Religious Insight*, 205. Emphasis removed from the original.

¹⁹. More precisely, that person is doing their moral and spiritual duty *as well as a finite being can*. On Royce’s view, no finite being can fully live up to the demands of morality. See Dwayne A. Tunstall, “Royce’s Ethical Insight: Its Relevance for Today”, Paper presented at Josiah Royce: Pragmatist, Philosopher of Religion, Ethicist, Hamburg, Germany, October 2, 2015.

cases.²⁰ Still, on Royce's view there truly are experiences of evil that no one would wish to remove from their lives.²¹

Whether they occur through conscious separation or accidental death, the evils that often cause the most psychological, emotional, and spiritual damage are the losses of loved ones.²² Supposing that one can forgive an unfaithful spouse or a friend-turned-enemy, that person's betrayal can never be forgotten. Even so, that grief is not something we would want to dispose of entirely. One might wish to numb oneself to the pain if it is unbearable, but not to the sensitivity or connection that causes the pain. For example, a person may wish to no longer grieve a departed loved one, but that person would certainly not wish to lose all the memories of their loved one, nor would they hope to never love another person again. Spiritual strength is acquired by recognizing and retaining social sensitivity through the grieving process and using it to deepen relations with others.²³ The result of that struggle is *sorrow*.

In order to recognize the strength that sorrow offers, we must take a step back from the pain of grief and recognize why it exists. A severed tie between intimately connected people gives rise to the worst imaginable pain. This implies that humans are profoundly social beings who have the capacity to intimately connect with others. Sorrow's unsettling prevalence presents a religious insight: spiritual strength is not won by merely avoiding possible suffering because, in this world, everyone will have sorrows. Neither one's world nor one's fellows are perfect. Individuals gain spiritual strength by developing the patience and courage to face a future full of meaningful relations without bitterness or resentment.

With this in mind, the next step is to endure the hardship. Finally, it is necessary to draw upon the insight of sorrow and reinvest oneself in the social reality. One must deepen relationships or form new ones while remaining aware that sorrow in the future is guaranteed.²⁴ There is no pain-free way to restore balance to life, but life can become much more meaningful in the process. By spinning grief into sorrow, one can enjoy new and renewed loyalties—the dedication to vital communities and causes that enrich one's life. However, this solemn work becomes especially difficult when one's grief is caused by betrayal.

When a person betrays their community, they set off a grieving process which is often fatal for the community. The losses incurred by the betrayal are permanent because the betrayer cannot undo the destructive deed. As much as the betrayer may wish they could turn back time, punishment only reminds them that their deed is irrevocable. Moreover, the community's memory of the evil deed scars any remaining affection or sympathy that the community can extend to the betrayer. However, as traumatic as it is, the aftermath of a betrayal is fertile ground for the creative power of communal action. Members who are willing

²⁰. Assuming, of course, that eliminating the remainder did not involve doing anything terribly imprudent.

²¹. Royce, *Sources of Religious Insight*, 239.

²². These cases of personal loss may be categorized as either moral or non-moral evils. An elderly grandmother dying in her sleep is obviously not committing a moral evil, since the grandmother did not choose to devastate her family. However, if that same grandmother was intentionally given a fatal dose of medication by a twisted attendant, her death would be an instance of moral evil.

²³. Royce, *Sources of Religious Insight*, 252.

²⁴. *Ibid.*, 253.

to bear the sorrow of betrayal and work to reestablish their community bring about goods that would have been impossible had the betrayal not taken place, by manifesting the spirit of their community.²⁵ To illustrate this point, we need to outline the process of atonement. In the *Problem of Christianity*, Royce illustrates his idea of atonement through an interpretation of the Biblical story of Joseph.²⁶ Joseph's brothers sold him into slavery because they were jealous of the preferential treatment he received from their father. Years later during a time of great famine Joseph's brothers travelled to Egypt, where Joseph served as Pharaoh's trusted advisor, to buy supplies. Joseph revealed his identity to his brothers and sent them back to their father with ample provisions. On Royce's view, when Joseph provided for his family he was engaged in a creative reversal of his brothers' betrayal or, in other words, an act of atonement. There are three central elements to such acts of atonement. The first is that the act is performed by some person other than the betrayer. The second is that the act is made possible by the specific betrayal for which it atones. The third element is that the act of atonement makes the world better than it was before the betrayal.²⁷ In this case, Joseph could not have helped his brothers if he were not sold as a slave. Joseph chose to see through his grief, endure it, and make it part of a process of reconciliation. Coincidence may have brought them together spatially, but only Joseph's actions could have reunited the family spiritually. Now that the foundation of Royce's views has been laid, we can make the divine thread running through the discussion explicit.

Recall that, for Royce, a community is a superhuman being. As such, communities can be afflicted by a kind of social blindness like the one that we discussed at the beginning of the section. When members set out to do things on behalf of their community, they are expressing love for one another and for that being that unites them: the spirit of their community.²⁸ However, the love for a community can, itself, become a stumbling block on the path toward creating more inclusive communities. I do not need to regale the reader with horrific stories from our species' past. Suffice it to say that people are in constant danger of allowing the love they have for their community to become obsessive and exclusionary.²⁹ When members refuse to recognize the value of external communities, their own communities stagnate, ossify, or turn malignant. This is social blindness scaled up to the communal level. The members mistake their finite, fallible community for the highest human good and do not seek to actualize any higher community. On the contrary, Royce's view is that the greatest good is the struggle to actualize the *highest* community, i.e., the Universal Community.

The Universal Community is precisely that "spiritual unity of all the world of reasonable beings" mentioned above.³⁰ Finite communities and their members work toward actualizing this ideal by guarding against encroaching blindness and remaining inclusive, uplifting, and

²⁵. Royce, *Problem of Christianity*, 180.

²⁶. Ibid. 204.

²⁷. Ibid. 180.

²⁸. Ibid. 265.

²⁹. Kegley, *Josiah Royce in Focus*, 93.

³⁰. Royce, *Sources of Religious Insight*, 205. Emphasis removed from the original.

faithful to the Spirit of the Universal Community, which guides the community-building process. The Spirit of the Universal Community, which is Royce's quasi-analogue to the Abrahamic God, is the divine being who calls upon individuals to conquer evil by exercising their loyal devotion to communities—especially through atoning deeds—and calls upon communities to seek common ground and solidarity with each other.

The community-building process is also the “evil-overcoming” process because community-building harmonizes the wills of rational beings and unites them under common causes. This means that all individuals and communities are tasked with the temporal, yet endless, process of overcoming evil. The task is temporal because it takes place within the processes of the world as the Spirit overcomes evil step by step through the triumphs of its members. The task is endless because, while individuals can work toward actualizing the ideal of the Universal Community, they will necessarily fail. As long as there are finite beings there will be inhospitable conditions and conflict.³¹ Even in principle, we cannot atone for every instance of evil. However, those who *act toward* that ideal state of atonement, and strive to bring lasting peace to this fallen world, take up the Spirit's cause as their own, and thereby find their place in the divine life.³²

Readers who are familiar with the argument from evil and the typical rebuttals it draws will wonder what Royce's answer has to offer to the discussion. In the next section I will offer a preliminary answer to that question.

III. Two Traditional Answers to the Problem of Evil

In this section I will argue that Royce's answer to the problem of evil has distinct advantages over two traditional theodicies because it does not presuppose the traditional conception of God. To make the advantages explicit, however, we will need to acquaint ourselves with those two traditional theodicies: soul-making and free will.

Soul-making theodicies presume that human spiritual development, culminating in the achievement of a spiritual ideal ordained by God, is supremely valuable. In fact, they presume that God created human beings for the express purpose of attaining that ideal and earning the right to dwell with God.³³ That being said, spiritual development comes at a price. People must endure evil in order to acquire the character traits necessary to develop according to God's plan. Since God created a world where people can develop through their struggles with evil and—at least potentially—achieve the spiritual ideal set out for them, God remains morally perfect. With this understanding, the existence of God is consistent with the existence of evil. Now we move to free will theodicies.

Free will theodicies presume that libertarian free will, when it is used to worship God and when it is in accordance with God's moral dictates, is supremely valuable. These theodices

³¹. Indeed, there could be no *finite* beings at all without there also being some conditions which undermine their existence.

³². Royce, *Problem of Christianity*, 186.

³³. René Van Woudenberg, “Chapter 12: A Brief History of Theodicy,” in *The Blackwell Companion to the Problem of Evil*, ed. Justin P. McBrayer and Daniel Howard-Snyder (Somerset, NJ.: John Wiley & Sons, Incorporated 2013), 177.

posit that God created people with free will so that they could worship and act morally of their own accord.³⁴ Although people may misuse their free will and act immorally, the great value of its proper use more than justifies the existence of evil. Therefore, God must have created people with free will in order to create a morally perfect world and the existence of God is not inconsistent with the existence of evil.

Royce would argue that on both views, God is responsible for the existence of unnecessary evils.³⁵ An all-loving God who is responsible for necessary evils can remain all-loving, but a God who allows *unnecessary* evils to transpire cannot be considered all-loving. Therefore, the theodicies are inadequate answers to the narrow problem of evil. To delineate these unnecessary evils, we will begin with the soul-making account. On this view, God is responsible for the suffering required by the developmental process as it exists now. It may very well be the case that persons develop spiritually by struggling with evil and eventually earn the right to dwell with God. But if that is the case, it is only so because God designed the world in such a way that suffering is necessary for spiritual growth. All other things being equal, a world where people do not have to suffer to acquire godly character traits is better than one in which they must. Any attempt to justify the soul-making process by appealing to the goods afforded by that process fails because

[T]alk of medicinal and disciplinary evil, perfectly fair when applied to our poor fate-bound human surgeons, judges, jailors, or teachers, becomes cruelly, even cynically trivial when applied to explain the ways of a God who is to choose, not only the physical means to an end, but the very *Physis* itself in which path and goal are to exist together.³⁶

Being omnipotent, God could have just as easily designed a soul-making process that did not involve the experience of suffering but chose not to do so. Now, we consider the free will account.

If someone watched a man stumble out of a bar, fumble with his keys, and proceed to drive away clearly intoxicated, we would hold that person accountable for not intervening if they were able to do so. Of course, an omniscient, omnipresent, and omnipotent God is perfectly able to intervene whenever, wherever, and in whatever fashion that God desires. So even if the value of free will is granted, God could ensure that innocent people do not die when drunk people take the wheel. Driving under the influence need not cause anyone but the driver and, perhaps his willing riders, to suffer. Yet, there are many innocent victims every year. The theodicies propose that God tests the innocent by letting them suffer at the hands of the guilty and purifies their souls through pain. Those are not expressions of omnibenevolence. Because of this, neither a free will theodicy nor a soul-making theodicy can rescue the traditional conception of God from the narrow problem of evil.

³⁴. Van Woudenberg, "Brief History of Theodicy", 185.

³⁵. Josiah Royce, "The Problem of Job", in *Studies in Good and Evil*, (New York: D. Appleton and Company, 1898). See also: Kegley, *Josiah Royce in Focus*, 79-80.

³⁶. Royce, "The Problem of Job", 9.

Unlike the traditional conception of God, the Spirit of the Universal Community does not create the world, but expresses itself through the existing processes of the world.³⁷ This means that the Spirit cannot conceivably bear responsibility for the existence of evil. Furthermore, while the traditional God allows evil to exist to suit its own salvific tastes, the entire aim of the Spirit is to overcome evil by bringing finite beings together in community.

A full comparison of Royce's answer to the traditional answers will have to wait for another work. However, I have shown in this work that Royce's answer is in no way deflationary or dismissive of the problem of evil, so long as the problem is properly understood. Royce does not concede that "evil" is, say, a non-cognitivist ascription in the way some philosophers have tried to argue that "murder is wrong" can translate to "boo murder!" Nor is evil reducible to a naturalistic truism such as "evil is whatever *homo sapiens* do not prefer in their environments or condone in the behavior of their fellows." Royce provides a teleological answer to the problem of evil that has clear advantages over traditional theodicies, but which does not suffer from the attendant difficulties of those theodicies. His answer merits scholarly attention because it poses a serious challenge to the Abrahamic status quo in the philosophy of religion.

IV. Conclusion

In this paper, I introduced the narrow problem of evil and Josiah Royce's answer to the broad problem of evil. After I reconstructed Royce's answer, I gave some preliminary reasoning as to why two traditional answers to the narrow problem of evil fail. Then I showed that Royce's conception of God does not face even a *prima facie* existential threat from the existence of evil, and its sole mission is to bring about the resolution of evil by bringing finite beings into community. Thus, Royce's answer is theistic and teleological, but it does not suffer from the deficiencies found in some traditional answers to the problem of evil. For these reasons, and the fact that answers to the narrow problem of evil are fraught with difficulties, I argue that Royce's answer to the broad problem of evil merits a response from philosophers of religion in the Abrahamic traditions.

³⁶. See Kegley, *Josiah Royce in Focus*, 157-8. For a brief summary of Royce's views on the monotheistic doctrine of creation.

References

- Foust, Matthew A. "Confess Your Contradictions: Schelling, Royce, and the Art of Atonement." *Journal of Speculative Philosophy* 26, no. 3 (2012): 516-530.
- Garchar, Kim. "Sin, Sorrow, and Suffering: A Roycean Response to These Deeper Tragedies of Life." *American Journal of Theology & Philosophy*, vol. 33, no. 1 (January 2012): 57-73.
- Kegley, Jacquelyn Ann K. *Josiah Royce in Focus*. Bloomington, Ind.: Indiana University Press, 2008.
- . "Josiah Royce." *Internet Encyclopedia of Philosophy*. Accessed July 15, 2015. <http://www.iep.utm.edu/roycejos/>.
- Oppenheim, Frank M. *Royce's Mature Philosophy of Religion*. Notre Dame, Ind.: University of Notre Dame Press, 1987.
- . *Reverence for the Relations of Life: Reimagining Pragmatism via Josiah Royce's Interaction with Peirce, James, and Dewey*. Notre Dame, Ind.: University of Notre Dame Press, 2005.
- Parker, Kelly A. "Josiah Royce." In *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition). Edited by Edward N. Zalta. Accessed July 15, 2015. <http://plato.stanford.edu/archives/sum2014/entries/royce>.
- Royce, Josiah. *The Sources of Religious Insight*. New York: C. Scribner's Sons, 1912. Reprint. Washington D.C.: Catholic University of America Press, 2001.
- . *The Problem of Christianity*. Chicago: The MacMillan Company, 1913. Reprint. Washington D.C.: Catholic University of America Press, 2001.
- . "The Problem of Job", in *Studies of Good and Evil*. D. Appleton and Company, New York: 1898.
- . *The Philosophy of Loyalty*. New York: The MacMillan Company, 1908. Accessed online July 25, 2015. <https://goo.gl/G5mzA4>.
- Smith, John E. "Royce's Spirituality: The Integration of Philosophical Reflection and Religious Insight." Introduction to *Josiah Royce: Selected Writings*. Edited by John E. Smith and William Kluback. Mahwah, NJ: Paulist Press, 1988. Accessed online July 24, 2015. <https://goo.gl/J1kQ7u>.
- Tooley, Michael. "The Problem of Evil." in *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition). Edited by Edward N. Zalta. Accessed July 16, 2015. <http://plato.stanford.edu/archives/sum2015/entries/evil/>.
- Van Woudenberg, René. "A Brief History of Theodicy," in *The Blackwell Companion to the Problem of Evil*. Edited by Justin P. McBrayer and Daniel Howard-Snyder. Somerset, NJ.: John Wiley & Sons, Incorporated 2013. ProQuest.

MINDS AND BODIES: EARLY MODERN SOCIAL JUSTICE

Clare Áine Keefer, Eastern Michigan University

Abstract

In *A Serious Proposal to the Ladies* parts I and II, Mary Astell argues that social conditioning impacts women's self-image in such a way as to prevent them from striving for scholarly achievement. Astell's solution is to allow women to withdraw from society into dedicated schools for women and by women, as an alternative to marriage and family life. In this paper, I will explore some of the implications of that argument, how it might be expanded to other marginalized populations, and argue that despite Astell's proposed solution being proven to create at least as many problems as it solves, the groundwork laid in her arguments can form a basis for a functional model of educational justice today. We have learned that "separate, but equal" education is not a solution to the problem of "achievement gaps" between privileged and marginalized populations. If social conditioning impacts educational drive and achievement for women, then it also impacts other oppressed populations. I maintain that subverting this structural oppression is a key to dismantling it and achieving educational justice. The typical foundations for educational justice come from the imperative that education makes better citizens or that education allows further education on a topic. I maintain that if the goal is educational justice, it is necessary to overcome the determinants of social conditioning.

Minds and Bodies: Early Modern Social Justice

Introduction

René Descartes is frequently cited as the "father of modern philosophy". His impact upon the field is challenging to overstate as there are still philosophers wrestling with the questions he raised some four centuries ago. Much less well known is Mary Astell, said to be the "first English feminist" (Batchelor, 2002). Though she was not progressive by today's standards, being a royalist and both politically and religiously conservative, still she wrote and published in the late seventeenth and early eighteenth centuries on the topic of women's intellectual advancement and equality of reason. While Descartes himself expressed a belief in equality between the sexes, his concept of mind/body dualism has often been cited by scholars as contributory to the systemic, institutional oppression of women, on the basis of their being

deemed less rational, and more closely affiliated with the body and its sensual nature than men. Much scholarship has been done on this topic, and a full exploration of that connection is outside the scope of this paper, but a brief survey will be included for context. The reader is encouraged to examine more closely the works referenced for a detailed treatment of that topic. Astell presented a philosophical account of minds and bodies that differed significantly from that offered by Descartes, holding that humans are a *union* of mind and body, and that one can act upon the other. She posits a metaphysics of differentiated individuals, each with unique abilities, wherein experience directly impacts an individual's ability to reason. It is upon this ground that she builds the argument that women are no less capable of reason than men, but that social conditioning leads them to believe – and thus act as if – they are. In this paper, I will explore Astell's arguments in contrast with those of Descartes to explicate this early modern example of a social justice orientated epistemology and some of its implications, as well as to explore the applicability of her concepts today.

Mind Body Dualism

In his *Discourse on the Method*, René Descartes clearly and definitively established mind/body dualism: “this me — this soul that makes me what I am — is entirely distinct from the body, is easier to know than the body, and would still be just what it is even if the body didn't exist” (Descartes, p. 15). In the culmination of *Meditations on First Philosophy*, he revisited this, writing:

My essence consists entirely in my being a thinking thing. [...] because on the one hand I have a clear and distinct idea of myself, insofar as I am merely a thinking thing and not an extended thing, and because on the other hand I have a distinct idea of a body, insofar as it is merely an extended thing and not a thinking thing, it is certain that I am really distinct from my body, and can exist without it. (p. 51)

Though mind/body dualism was hardly a new way of thinking about being, Descartes' formulation of the idea came to be broadly embraced. A positive result of this was the furtherance of medicine. A doctrine of the Catholic Church called “the resurrection of the body” maintains that for a person to go to Heaven at the Resurrection, the body needs to be intact. As a result of this doctrine, many countries prohibited or severely limited the study of anatomy through dissection. Dr. Neeta Mehta, in her article, “Mind-body Dualism: A Critique from a Health Perspective”, writes: “[T]here was a religious prohibition on the study of human anatomy through dissection. Descartes, through mind-body dualism, demythologized body and handed over its study to medicine” (Mehta, 2011). While this is certainly a positive outcome, not every use of Descartes' work was so benign. As Elizabeth V. Spelman writes:

[W]hen one recalls that the Western philosophical tradition has not been noted for its celebration of the body, and that women's nature and women's lives have long been associated with the body and bodily functions, then a question is suggested. What connection might there be between attitudes toward the body and attitudes toward women? (1982, p. 110)

Dr. Sarah E. Johnson has an answer to Spelman's question: "As allegedly more subject to the body than men were, women possessed less reason and lacked control over their passions, determined in large part, of course, by the body's humoural balance" (p. 12). Johnson goes on to explain that "[n]ot only were women more bodily than men, but their bodies were also physiologically inferior [...] matching their easily swayed mental constitutions" (2014, p. 12). Whilst this misogyny did not originate with Descartes, his formulation of mind/body dualism at the very least contributed to the justifications for the attitude toward women, and thus to the oppression engaged in under such justification. Genevieve Lloyd provides a stark example, quoting Hegel:

Women are educated – who knows how? – as it were by breathing in ideas, by living rather than by acquiring knowledge. The status of manhood, on the other hand, is attained only by the stress of thought and much technical exertion. (qtd. in Lloyd, 38)

Lloyd adds, "In western thought, maleness has been seen as itself an achievement, attained by breaking away from the more 'natural' condition of women" (1984, p. 38). This is not to say that these justifications have ever been required in human history to fuel prejudice against any marginalized group – in fact, it is often that these arguments are posed as a means of legitimizing already extant prejudices. Just as with the prejudice against Black people in the writings of the early modern period of philosophy – as explored by many critical race theorists, including Emmanuel Eze, Barbara Hall, and Debra Nails, just to name a few – misogyny has a long history of pseudo-scientific and ostensibly "rational" arguments made by predominantly white men to attach a veneer of reason to a markedly unenlightened emotional reaction to perceived threats to their hegemony. The very concept of "reason" carries a bias toward white, European masculinity.

Equality and Gendering of Reason

We can look back as far as Pythagoras and see "female" equated with "bad", and "male" with "good" in his famous table of opposites. Phyllis Rooney, in "Gendered Reason: Sex Metaphor and Conceptions of Reason," writes that, with the rise of Enlightenment ideals, many of the images of male as dominant, causal, and active and female as incomplete, in subjugation, and passive may appear to be left behind, "but we get what is at best a shift in the articulated, explicit claims about reason and mind" (p. 82). She goes on to note that Descartes "allowed that 'even women' could develop his rational method." In spite of Descartes' generosity toward women, Susan Bordo finds in Descartes a culmination of a "rebirthing of nature (as machine) and knowledge (as objectivity)", resulting in a "supermasculinized model of knowledge in which detachment, clarity, and transcendence of the body are all key requirements" (p. 50). Genevieve Lloyd delves into this concept in her book, *Man of Reason: "Male" and "Female" in Western Philosophy*. She writes, "through [Descartes'] philosophy, Reason took on special associations with the realm of pure thought, which provides the foundations of science, and with the deductive ratiocination which was of the essence of his method" (p. 49). She continues:

We owe to Descartes an influential and pervasive theory of mind, which provides support for a powerful version of the sexual division of mental labour. Women have been assigned responsibility for that realm of the sensuous which the Cartesian Man of Reason must transcend, if he is to have true knowledge of things. (1984, p. 50)

This “Cartesian Man of Reason” is one who has sufficient free time to spend long periods in meditation, who holds fast to reason (the realm of the mind), as opposed to non-reason (the realm of the body), he not only judges, but specifically judges *well*, and, as Descartes enumerates in his *Discourse on Method*, he is capable of “telling the true from the false”; he is able to learn anything just as well as anyone else if they simply take their thoughts along the appropriate path (p. 1). Though Descartes held that “even women” could develop his rational method, common sentiment of the day held that women “are naturally incapable of acting prudently” and “necessarily determined to folly” (Astell, 2014). Here is enshrined the binary between mind and body, and thus, between men and women.

Mind/Body Unions

In contrast to Descartes’ dualism, Mary Astell asserts that people are unions of minds and bodies. In *A Serious Proposal to the Ladies*, she writes: “We know and feel the Union between our Soul and Body, but who amongst us sees so clearly, as to find out with Certitude and Exactness, the secret ties which unite two such different Substances, or how they are able to act upon each other” (Astell, 1994, 1994, p. 101)?

If all people are unions of both mind and body, then the marginalization of women on the basis of their ‘natural condition’ is revealed as problematic. Rather than being creatures of ‘pure reason’, men are just as ‘bodily’ as women, just as susceptible to passions, tempers, and the other aspects of their physicality as impedances to their reason. It is upon this ground that Astell builds her argument that women are no less capable of reason than men. But if this is so, how do we explain the gross disproportionality of educational achievement between men and women during the early modern period? According to Astell, it is nothing more than the natural result of the oppression resulting from the belief in mind/body dualism, and the nature of women as “more bodily than men.”

Social Conditioning and its Effects upon Reason

Astell argues that social conditioning is the cause of an achievement gap between women and men, educationally. In responding to the accusation that “women are naturally incapable of acting prudently or that they are necessarily determined to folly” she writes:

The incapacity, if there be any, is acquired, not natural . . . The cause therefore of the defects we labor under is, if not wholly, yet at least in the first place, to be ascribed to the mistakes of our education which . . . spreads its ill influence through all our lives. (Astell, 2014, pp. 55-56)

She goes on to more clearly, and rather acerbically, cast the blame at the feet of patriarchal society:

Women are from their very infancy debarred those advantages with the want of which they are afterwards reproached and nursed up in those vices which will hereafter be upbraided to them, so partial are men as to expect brick where they afford no straw and so abundantly civil as to take care we should make good that obliging epithet of *ignorant* which out of an excess of good manners they are pleased to bestow on us! (2014, p. 56)

Her argument is that experiences, particularly experiences of oppression and degradation, train a person to believe that all they are capable of is that to which they are constantly being told they are limited. It is popular, particularly in educational circles, to sum this up as “students rise (or fall) to the level of expectations.” The majority of women in her day were uneducated and did not actively seek out education, not because they were incapable of being educated, but because society had conditioned them to believe that God created them as “lesser” and incapable. In such a circumstance, Astell asserts that women adopt the notion that they are constantly being told, that they are “naturally proud and vain,” and do not strive beyond that. This leads into a feedback loop, wherein women do not strive, thus they do not achieve, confirming the perception that they are incapable of achievement, and therefore, opportunities for self-improvement are “wasted” on them. This scenario can easily be seen in many times and places over human history, continuing to the present day, and women are by no means the only victims of such oppression.

Social Justice Implications of the Effects of Social Conditioning

This feedback loop happens to people occupying many other axes of oppression. We speak today of achievement gaps in inner city schools. We segregate differently abled students into “special education” classes, regardless of the nature or severity of their differentiation of ability. We had a Supreme Court Justice commenting about those who believe that “it does not benefit African-Americans to get them into [elite universities] where they do not do well, as opposed to having them go to a less-advanced school, [...] where they do well” (Fisher v. UT Austin, 2015). Astell’s solution was to establish separate schools for women, but United States history has clearly demonstrated the results of a “separate but equal” educational system. How, then, can this situation – which is obviously still a problem today – be resolved?

Socially Conscious Education

Any attempt to rectify the structural injustice built into educational models needs to address both sides of this issue. José Medina, in his book, *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*, asserts that the experience of being oppressed can, in some ways, present an epistemic advantage, in that oppression provokes the formation of learning processes that the privileged have less opportunity to develop. The oppressed, he writes, “tend to be better listeners, having a *more acute attentiveness to hermeneutical gaps*” (Medina, 2012). Oppressed agents see more of the oppressive structures than privileged agents, as they are the ones who keenly feel the effects of those structures. Privileged agents often struggle to see institutional oppression in the structures of society, much less the ways in which they are participatory. This is why, if we wish to dismantle structural oppression and create a just educational system, education must

actively affirm the lucidity and capabilities of vulnerable agents and work to counter the social conditioning these agents often have to overcome. Further, such a system must also actively engage in educating the ignorance of privileged agents on the topic of systemic oppression. Educational programs must, at all levels, both reflect and teach diversity. The study of cultures and people occupying axes of oppression must cease to be segregated into various tracks of cultural studies, and, instead, be included as part of core curricula at all levels. If we would live up to Descartes' belief in the equality of ability, we must follow in the steps of Mary Astell, and accept that the differences between us are just that: *differences*, not limitations. Different bodies are not *lesser* bodies, and neither are the minds or capacities to achieve associated therewith.

References

- Uebel, M. (1996). *Unthinking the Monster: Twelfth Century Responses to Sacrean Alterity*. In J. Cohen, *Monster Theory: Reading Culture*. Minneapolis: university of Minnesota Press.
- Wright, A. (2013). *Monstrosity : The Human Monster in Visual Culture*. London: I.B. Tauris & Co LTD.
- Asma, S. T. (2009). *On Monsters: An Unnatural History of Our Worst Fears*. Oxford: Oxford University Press.
- Astell, M. (2014). *A Serious Proposal to the Ladies*. (S. L. Jansen, Ed.) Saltar's Point Press.
- Astell, M. (1994). Selections from *A Serious Proposal to the Ladies, Part II*. In M. Atherton (Ed.), *Women Philosophers of the Early Modern Period* (pp. 99-125). Indianapolis: Hackett Publishing Company, Inc.
- Batchelor, J. (2002, March 21). *Mary Astell*. Retrieved November 7, 2015, from Literary Encyclopedia.
- Benhabib, S. (1994). In Defense of Universalism. Yet Again! A Response to Critics of Situating the Self. *New German Critique* .
- Bordo, S. (1999). Selections from *The Flight to Objectivity*. In *Feminist Interpretations of René Descartes* (pp. 48-69). University Park, PA: The Pennsylvania State University Press.
- Brown University . (n.d.). *Liber Monstrum: The Book of Monsters of the Sea*. Retrieved October 7, 2014
- Cohen, J. (1996). *Monster Theory: Reading Culture*. Minneapolis: university of Minnesota Printing Press.
- Creed, B. (1994). *The Monstrous Feminine: Film, Feminism, and Pyschoanalysis* . New York: Routledge.
- Eagleton, T. (2014). The Body as Langauge. *Canadian Review of Comparative Literature* , 11-16.
- Descartes, R. (2005, June). *Discourse on the Method of Rightly Conducting one's Reason and Seeking Truth in the Sciences*. (J. Bennett, Trans.)
- Descartes, R. (1993). *Meditations on First Philosophy*. (D. A. Cress, Trans.) Indianapolis: Hackett Publishing Company, Inc.
- Fisher v. University of Texas at Austin, et al, 14-981 (Supreme Court of the United States December 9, 2015).
- Foucault, M. (1999). *Abnormal: Lectures at the College De France*. New York: Picador.
- Johnson, S. E. (2014). *Staging Women and the Soul-Body Dynamic in Early Modern England*. Surrey: Ashgate Publishing Limited.
- Kimmel, M. (2002). *Toward a Pedagogy of the Oppressor* . *Tikkun Magazine* .
- Kristeva, J. (1982). *Powers of Horror: An Essay on Abjection*. New York: Columbia university Press.

- Lloyd, G. (1984). *Man of Reason: "Male" and "Female" in Western Philosophy*. Minneapolis: University of Minnesota Press.
- Nuzzo, L. (2013). Foucault and the Engima of the Monster. *Int J Semiot Law* .
- Nanda, S. (1996). *Hijra*. Retrieved November 4, 2014
- Martins, H. (2010). Wittgenstein, the body, its metaphors. *DELTA* .
- Medina, J. (2012). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. Oxford University Press.
- Mehta, N. (2011). Mind-body Dualism: A critique from a Health Perspective. *Mens Sana Monographs* , 9 (1), 202-209.
- Pender, S. (1996). No Monsters at Resurrection. In J. Cohen, *Monster Theory*. Minneapolis: the university of Minnesota Press .
- Symbolic and Structural Approaches to Place*. (n.d.). Retrieved Novemeber 23, 2014, from Pegas Server-univeristy of Central Florida: <http://pegasus.cc.ucf.edu/~janzb/courses/hum3930b/foucault1.htm>
- Spelman, E. V. (1982, Spring). Woman as Body: Ancient and Contemporary Views. *Feminist Studies* , 8 (1), pp. 109-131.
- Richter, D. (n.d.). *Ludwig Wittgenstein (1889—1951)*. Retrieved November 12, 2014, from Internet Encylopedia of Philosophy.
- Rooney, P. (1991). Gendered Reason: Sex Metaphor and Conceptions of Reason. *Hypatia* , 6 (2), 77-103.

PREDICTING THE TRUTH: OVERCOMING PROBLEMS WITH POPPER'S VERISIMILITUDE THROUGH MODEL SELECTION THEORY

K Raleigh Hanson, Washington State University

Abstract

The purpose of this research is to investigate the possibility of using aspects of model selection theory to overcome both a logical problem and an epistemic problem that prevents progress towards the truth being measured while maintaining a realist approach to science. Karl Popper began such an investigation into the problem of progress in 1963 with the idea of verisimilitude, but his attempts failed to meet his own criteria, the logical and epistemic problems, for a metric of progress. Although philosophers have attempted to fix Popper's verisimilitude, none have seemed to overcome both criteria yet. My research analyzes the similarities between Predictive Accuracy (PA) and Akaike's Information Criterion (AIC), both parts of model selection theory, and Popper's criteria for progress. I find that, in ideal data situations, it seems that PA and AIC satisfy both criteria; however, in non-ideal data situations, there are issues that appear. These issues present an interesting dilemma for scientific progress if it turns out that our theories are in non-ideal data situations, yet PA and AIC seem to be better overall indicators of scientific progress towards the truth than other attempts at overcoming the problems of Popper's verisimilitude.

Predicting the Truth: Overcoming Problems with Popper's Verisimilitude Through Model Selection Criteria

One problem when discussing scientific progress is whether or not our current theories have made any progress towards the truth, or have just become better predictive tools. There is an intuitive notion that newer theories are truer than older theories because they appear to identify more true causes of a target system. However, it turns out that it is notoriously difficult to provide an analysis of what it means for one theory to be closer to the truth than another theory. The issue is even more pronounced when considering the pessimistic meta-induction: since all of our past theories have been false, it is likely that all of our current theories will also be false and our future theories as well. This poses a problem for scientific realism which holds that identifying the true causes of a target system is an important aim of science.

While the discovery of new causes that affect target systems does seem to be an important part of scientific progress, it is not clear that increasing the ability to predict the behavior of target systems will always correspond to knowing more causes of that system (Forster and Sober 1994). In fact there is some evidence that our best predictive models and theories might not always be our best explanatory models and theories (Goldsby 2013). However, if we want to define progress in realist terms, there needs to be some account of what proximity to the truth is and how newer theories get us closer to the truth. I will refer to these two concerns as the logical problem and the epistemic problem respectively.

An early attempt to overcome the logical and epistemic problems was introduced by Karl Popper in his work *Conjectures and Refutations*. Popper (1963) called his attempt to overcome the two problems verisimilitude. The concept behind verisimilitude is intuitive in nature – a theory is closer to the truth if it makes more true claims and fewer false claims – but his later commentators would point out critical flaws such that verisimilitude was found to be inadequate for solving either the logical or the epistemic problem. A number of attempts have been made to revise or fix Popper’s language to make verisimilitude work, but none have overcome both the logical and epistemic problems. However, if progress can be defined as overcoming the logical and epistemic problems, then it is possible there may exist a framework elsewhere that satisfies that criteria.

One possible framework, predictive accuracy (PA), is a measure of the ability of a model to predict new data given old data. One plausible assumption is that the true model will be maximally predictively accurate, so increasing predictive accuracy will get one closer to the truth. According to Forster and Sober (1994), PA may be estimated using Akaike’s Information Criterion (AIC). If PA can be a measure of closeness to the truth, then using a model selection framework like AIC can select models closer to the truth. If, in turn, AIC can select a model that is closer to the truth because it is more predictively accurate than competing models, AIC can be useful for estimating progress. In this way, PA overcomes the logical problem by being a measure of how one model can be closer to the truth than another, and AIC overcomes the epistemic problem by showing that, when a new model is selected, it is because of both its increased proximity to the truth as well as its ability to predict new data.

The main concern for this paper is to investigate whether PA and AIC actually can overcome the logical and epistemic problems. I will begin by explaining why a notion of verisimilitude is important for the progress of science. I will then provide some background to Popper’s account of verisimilitude, and I will introduce model selection theory and explain how PA and AIC appear to satisfy the criteria demanded by verisimilitude. I will argue that PA and AIC can overcome both problems while in an ideal data situation and discuss what may occur while in non-ideal data situations. Finally, I will address the problems of PA and AIC as a form of verisimilitude and discuss what sort of progress we may actually have made.

Why is Progress Towards the Truth Important?

There are two basic accounts of the goals of scientific inquiry: realism and instrumentalism. Scientific realism maintains a concern for understanding the truth behind phenomena including things that can’t be directly observed. Even if the pessimistic induction is right, realism holds that newer theories can be closer to the truth than older theories. For

example, it seems correct to say that even though Copernicus's heliocentric model of the solar system is false, it is still closer to the truth than Ptolemy's geocentric model.

Unlike realists, instrumentalists view scientific theories as tools that help capture or predict observable phenomena regardless of the truth-value of the theories themselves (Chakravartty 2014). In this way, an instrumentalist values theories that can predict or account for observable phenomena even if we can't know the truth about the unobservable commitments of that theory (Van Fraassen 1980). Instrumentalists believe that the truth of unobservables is inaccessible and science should be aimed at predicting observable phenomena rather than identifying all and only true causes.

Although Popper was a realist, his critics would point out that his hypothetico-deductive approach³⁸ to science by falsifying theories only winnows away at an infinite set of false theories and this does not constitute actual progress. Popper's (1963) verisimilitude was his attempt to show that false theories could have degrees of closeness to the truth, and that removing false theories does constitute progress towards the truth. Popper hoped that verisimilitude would allow him to be a realist while still holding to his hypothetico-deductive approach to scientific inquiry. If progress towards the truth is the goal of science as Popper claims, then discarding an instrumentalist approach is an important step.

Popper's Verisimilitude

Popper correctly identified the logical and epistemic problems that must be overcome for verisimilitude to provide a measure of progress. The aims of verisimilitude can be easily formulated as the following questions:

- (A) Can we explain how one theory can be closer to the truth, or has greater verisimilitude than another?
- (B) Can we show that scientific practice has sometimes led to theories which are closer to the truth than their predecessors? (Forster; ms)³⁹

The first question addresses the logical problem: we must have an account of how one theory is closer to the truth than another. The second question addresses the epistemic problem. Given our epistemic limitations, we must be able to determine that the selection of one theory over another is actually progress towards the truth.

Of course, Popper had to clarify how the degrees of truth would be measured. Popper's (1963) intuitive definition of verisimilitude, V_s , of theory A is based upon a measure of the true and false contents of A . The $Ct(A)$ is made of all of the logical consequences of A and can be divided into truth content, $Ct_T(A)$, and false content, $Ct_F(A)$. Truth content of A is the set of all claims that are true in $Ct(A)$, and false content is the set of all claims that are false in $Ct(A)$. $Ct_F(A)$, subtracted from $Ct_T(A)$ provides a measure of verisimilitude:

³⁸ Popper's (1959) hypothetico-deductive approach was presented in his *Logic of Scientific Discovery*. According to Popper's method, a hypothesis should be formed in a way that can be deductively falsified rather than supported by evidence.

³⁹ Forster credits an unpublished manuscript by Alan Musgrave for this formulation of the logical and epistemic problems.

$$Vs(A) = Ct_T(A) - Ct_F(A) \text{ (Popper 1963, 234)}$$

This intuitive definition provides the basic notion behind verisimilitude within a single theory by determining the number of true and false logical consequences of theory A . The intuitive notion behind this measure is simple; it provides a measure $Vs(A)$ based upon $Ct_T(A)$ and $Ct_F(A)$. By quantifying the true and false content of theories, this definition would allow two theories, A and B , to be compared as follows:

$$Vs(A) > Vs(B) \leftrightarrow [Ct_T(A) - Ct_F(A)] > [Ct_T(B) - Ct_F(B)]$$

The intuitive definition is a good first pass at the logical problem, but real theories are more complicated. For example, assume there are two theories, A and B , and that theory A and theory B are both false. To explain this concept, Popper (1963) offers the following example for any given theory: assume that today is Monday and theory A states that today is Tuesday; although theory A is false, it still entails true logical content such as today is not Wednesday and today is either Monday or Tuesday (Popper 1963). Because there are an infinite number of consequences, the Popper's first pass can't actually serve as a measure of verisimilitude.

Popper improved upon his first pass by using set-theoretic terms to create a contrastive definition of verisimilitude. Popper's (1963) contrastive verisimilitude (PCV) can be stated as follows:

$$(PCV) \quad Vs(A) < Vs(B) \leftrightarrow [Ct_T(A) \subset Ct_T(B)] \wedge [Ct_F(B) \subseteq Ct_F(A)]$$

That is to say that for B to have greater verisimilitude, B must make every true claim made by A and at least one additional true claim not made by A , and every false claim made by B must also be made by A without any additional false claims.

As an example, consider Ptolemaic astronomy and Copernican astronomy. For the sake of simplicity, suppose that the only difference in content between Ptolemaic astronomy and Copernican astronomy is the location of the sun and the Earth. Copernican astronomy makes one true claim not made by Ptolemaic astronomy, the Earth revolves around the sun. Ptolemaic astronomy makes one false claim not made by Copernican astronomy, the sun revolves around the Earth. If PCV holds, Copernican astronomy has greater verisimilitude because it makes all the true claims that Ptolemaic astronomy makes plus an additional true claim, all the false claims made by Ptolemaic astronomy are also made by Copernican astronomy, and Copernican astronomy makes one fewer false claim. Popper had examples like this in mind when he developed PCV to satisfy the criteria for verisimilitude.

The Problem with Popper's Verisimilitude

PCV, however, is also problematic in a similar manner to Popper's intuitive definition. Working independently, Pavel Tichý (1974) and David Miller (1974) both discovered a critical logical flaw to PCV. Tichý and Miller both pointed out that two competing false theories will never meet the subset relations PCV lays out because whenever a new true consequence is added, a new false consequence is added as well. Consider the following claims:

- P1: The sun revolves around the earth
 P2: The planets move in perfect circles
 C3: The Earth revolves around the sun

Of course, we now know that P1 and P2 are false and C3 is true. The Ptolemaic model says P1 and P2 are true. The Copernican model says P2 and C3 are true. Now consider the following claim:

C4: P2 and C3

C4 is false because any conjunction that contains one false conjunct is always false. It is also a false claim that is not contained within the Ptolemaic theory. This can be called the conjunction problem. A true claim made by theory *B* but not made by theory *A* can be conjoined with a false claim made by theory *B* to create a new false claim not made by theory *A*. Thus, PCV will fail.

The incomparability of false theories is one of the consequences that developed from analysis of Popper's theory of verisimilitude. Tichý's and Miller's treatments of Popper's work show that it is impossible to add true consequences to a theory without also adding false ones, and equally impossible to subtract false consequences without also subtracting true ones. Two theories cannot be compared in terms of scientific progress towards the truth as Popper has defined it either as an intuitive notion or through PCV.

Applying Model Selection as Verisimilitude

If the concern of verisimilitude is to produce results that show theory progression is moving towards the truth by overcoming the logical and epistemic problems, it may be possible to look to forms of model selection that could serve the same purpose. A model is simply a set of equations that contain a number of adjustable parameters that is used to explain or predict a phenomenon (Forster 2000). A model can be broken down into the following parts: parameters, variables, and error terms. Consider the following toy models:

$$\begin{aligned} (M1) \quad & y = ax_1 + e \\ (M2) \quad & y = ax_1 + bx_2 + e \\ (FIT) \quad & y = 7x_1 + 0 \end{aligned}$$

In the above models, y is the dependent variable, x_1 and x_2 are independent variables, a , and b are adjustable parameters, and e is an error term to correct for observational errors. FIT is a *fitted model* where all the parameters are fixed. M1 and M2 represent families of curves or fitted models. For example, M1 represents all the curves that could occur when values are applied to the parameters. Note that FIT is a member of the family of fitted models of M1 (and M2).⁴⁰ The dependent variable is the measurable quantity of interest, and the independent variables are the causes that influence that quantity.

Model selection is concerned with fitting models to data, a process called curve fitting. Consider a graph of collected data. From a realist perspective, it is assumed that there is a true

⁴⁰ All of the fitted models of M1 are within the family of M2 where b equals 0.

curve that generates the data (give or take observational error). The goal of model selection is to find a model that is as close to the true curve as possible given the available data. Practicing scientists know that when the data set is small, simpler models tend to be better predictors than more complex models. In fact, it is well known that curves that perfectly go through every data point tend to be poor predictors because they overfit the data. The problem with overfitting is that it mistakes observational error for a true cause of the target system. If the goal of scientific realism is to discover true causes, and model selection can be used to identify true causes of a target system by avoiding overfitting and increasing PA, it may be possible to use model selection criteria to overcome the problems of verisimilitude.

Predictive Accuracy and AIC

Predictive accuracy, as defined by Forster and Sober (1994), is the ability for a selected model to predict new data given existing data. In situations where there is little data available, a simple model may be more predictively accurate than a more complex one, but as more data becomes available, the choice of models may be revised because the simpler model fails to be as predictively accurate. For example, in data poor situations, a simple model like M1 may be more predictively accurate, but, as the amount of data increases, a more complex model like M2 may be selected because of its greater ability to predict new data.

Although there are many types of model selection theories, this paper is concerned with AIC due to its relation to verisimilitude. Forster (2000) explains that an important part of AIC is that “the conclusions of AIC are . . . about its closeness to the truth” (213). If the true curve is maximally predictively accurate, and if AIC chooses the maximally predictively accurate curve given the data available, increasing PA can overcome the logical problem and AIC should overcome the epistemic problem.

The purpose of AIC is to minimize the Kullback-Leibler distance⁴¹ (K-L) between potential fitted curves within a family and the true curve represented by the data (Forster 2000). K-L distance, as defined by Burnham and Anderson (2002), indicates the distance between a candidate model and the true curve. However, since K-L distance cannot be computed without a prior knowledge of the true curve, a selection criterion like AIC must be used (Burnham and Anderson 2002). AIC, then, is supposed to provide an estimation of the closeness to the truth of a model. Sober (2008) provides the following formulation of AIC:

$$AIC(M) =_{df} \text{Log}\{\text{Pr}[Data|L(M)]\} - k$$

In this formulation, $L(M)$ represents the likeliest fitted model of M given the data available. $AIC(M)$ is found by taking the log likelihood of $L(M)$ and subtracting a penalty for complexity, k . The term k represents the number of parameters in the model and is used to prevent AIC from overfitting a model given the data when models are being compared. Complex models always fit the data better than simpler models, but as noted earlier, complex models are not always better predictors due to problems of overfitting. By having the correction for complexity, AIC is

⁴¹ It is worth noting that the K-L distance is not a true distance because it does not satisfy the triangle inequality. However, for the purposes of this paper the term “distance” works to clearly relate the concept of closeness or proximity between curves.

able to provide a reliable estimate of the model's PA. Thus, AIC only selects a model with a greater number of parameters when the log likelihood overcomes the k penalty.

Because AIC scores are dependent on the size of the data set, as the amount of data increases, AIC could select more complex models. For example, assume that there are three candidate models:

$$\begin{aligned} (M1) \quad & y = ax_1 + e \\ (M2) \quad & y = ax_1 + bx_2 + e \\ (M3) \quad & y = ax_1 + bx_2 + cx_3 + e \end{aligned}$$

In a data poor situation, AIC might favor the simpler model such that the following inequality holds: $AIC(M1) > AIC(M2) > AIC(M3)$. As we gather more evidence and the size of the data set increases, the AIC might recommend $M2$ over $M1$ if the AIC score of $M2$ is greater than $M1$. If it is true that x_2 is a new cause affecting the system, then it may seem that increasing PA will likewise increase closeness to the truth. In this way, the use of PA and AIC makes great progress dealing with both the logical and epistemic problems. Forster and Sober (1994) indicate that minimizing K-L distance to the true curve is the same as maximizing predictive accuracy. When selecting a model with the best AIC score, the model being selected is the closest model to the true curve given the available data.

The contrastive nature of PA and AIC also seem to overcome the epistemic problem that PCV failed to do. As new data is gathered, AIC may select a different family of curves with greater predictive accuracy than the current model. Because there is an existing metric of truth with the AIC score, obtaining a better score and increasing PA provides a contrastive view of progress similar to what Popper had attempted to do with PCV. In the examples of $M1$, $M2$, and $M3$ above, when AIC selects $M2$ over $M1$, an increase in closeness to the truth is being made along with an increase in predictive accuracy. That is, the new model is capturing more true causes of the target system while increasing the ability to accurately predict new data.

When AIC Fails

However, the ability for PA and AIC to overcome the logical and epistemic problems is based on ideal data situation. In data poor or data rich situations, there are complications that arise and create interesting dilemmas. Assume, for example, there is a target system that has three causes previously identified; however, the size of the data set is small. Even though we may know there are three causes of the target system, AIC may select a simpler model with only one cause because it will have greater predictive accuracy instead of a model that includes all three causes and is closer to the truth. This wrinkle may seem minor, but it shows that AIC may be tracking our ability to predict new data rather than tracking a theory's closeness to the truth in such a way that, while it can overcome the logical problem, it only does so in ideal data situations. However, the epistemic problem is still answered since, as data increases, AIC selects models that do identify more true causes of the target system as the predictive accuracy increases for those models.

Before turning to the next dilemma, the e term for observational error must be discussed. All of our scientific inquiry is subject to observational error or noise that is included in a data set. AIC assumes that observational error is present and accounts for it, but the very

presence of observational error is what leads to a greater problem behind AIC. There is a possibility that AIC will fail in data rich situations by selecting models that are further from the truth. While the error term included in models is supposed to deal with observational errors, as data sets get larger, there is a chance that AIC will recommend an additional parameter that is not a cause of the system being investigated. In other words, our model selection framework might be tracking the cause of observational error and mistakenly attributing it as a cause of the system under investigation. Forster and Sober (1994) explain that AIC was designed to estimate the size of the overfitting factor, but they also mention that the process is fallible. Given the possibility for AIC to recommend an error term as a new cause, we are now left with an interesting dilemma wherein either the logical problem or the epistemic problem will reassert itself. I will consider each horn of the dilemma separately.

I will begin by addressing the first horn. If our goal is to discover all the true causes affecting the target system, then in data rich situations we cannot be sure that a newly discovered variable is representing a cause of the target system or a cause of our observational error. If AIC is identifying causes of something outside of the target system, then there are some cases where we cannot tell whether progress is being made even if we are increasing predictive accuracy.

To illustrate the second horn of the dilemma, we can consider how a defender of the model selection framework might reply to the first horn. One might maintain that increasing PA always gets us closer to *some* truth. However, the truth being identified by increasing PA ceases to be about the target system, but begins to track the truth about the system that generates the data. This new system would take account of both the target system *and* the causes of our observational error. In such a situation, we give up the noumena in favor of the phenomena – we exchange our realist notion of the truth of a target system for the appearance created by the data. It is hard to see how such a solution would be palatable to scientific realists. Since the logical problem was supposed to allow for scientific realism, it seems that such a step gives up on the logical problem altogether.

These two horns of AIC create a trade off when dealing with the logical and epistemic problems. Either we accept that our choice in models can select better theories but we cannot always tell if we are getting closer to the truth, or we give up on scientific realism in favor of the notion that models with greater PA are closer to the truth about the system that gives rise to the data but not the true target of our inquiry.

Conclusion

PA and AIC seem to be heading in the right direction in understanding progress. However, if providing answers to Popper's logical and epistemic questions are the criteria by which a true sense of progress can be determined, PA and AIC seem to fall short of the mark if we want to maintain a realist approach to progress in all cases. The problem of data poor situations can be overcome by increasing the size of the data pool, and progress towards the truth can still be made. However, in data rich situations that may not be the case. Although AIC runs into this problem at the extreme limit, and it's likely that our extant theories have yet to run into it, there is a possibility that AIC will stop modeling the true causes of the target system at some point, and increasing PA will no longer be progress towards the truth of the target system. Of course, increasing PA and selecting a model with the best AIC, in ideal data

situations, does seem to satisfy both the logical and epistemic problem, so it may give progress hope. In terms of theories that can capture closeness to the truth and the movement of progress, PA and AIC seem to come closer than Popper's first attempt. Reminiscent of Popper's hypothetico-deductive method, PA and AIC seem to hold up to more severe tests than Popper's theory of verisimilitude did, and, in some ways, that seems like it is progress in itself.

References

- Burnham, Kenneth P. and David R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York: Springer, 2002.
- Chakravartty, Anjan. "Scientific Realism." *The Stanford Encyclopedia of Philosophy*. Last modified 2015. <http://plato.stanford.edu/archives/fall2015/entries/scientific-realism>.
- Forster, Malcolm. "Key Concepts in Model Selection: Performance and Generalizability." *Journal of Mathematical Psychology* 44 (2000): 205-231, <http://philosophy.wisc.edu/forster/papers/JMP2000.pdf>.
- Forster, Malcolm. "Verisimilitude and Likelihood." Manuscript, University of Wisconsin: Madison, 2004.
- Forster, Malcolm and Elliot Sober. "How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions." *The British Journal for the Philosophy of Science* 45, No. 1 (1994): 1- 35, <http://www.jstor.org/stable/687960>.
- Goldsby, Michael. "The "Structure" of the "Strategy": Looking at the Matthewson-Weisberg Trade-off and Its Justificatory Role for the Multiple-Models Approach." *Philosophy of Science* 89, no. 5 (2013): 862-873, <http://www.jstor.org/stable/10.1086/673728>. .
- Miller, David. "Popper's Qualitative Theory of Verisimilitude." *The British Journal for the Philosophy of Science* 25, No. 2 (1974): 166-177, <http://www.jstor.org/stable/686821>.
- Oddie, Graham. "Truthlikeness." *The Stanford Encyclopedia of Philosophy*. Last modified 2014. <http://plato.stanford.edu/entries/truthlikeness>.
- Popper, Karl. *Conjectures and Refutations*. London: Routledge, 1963.
- Popper, Karl. *The Logic of Scientific Discovery*. New York: Routledge, 1959.
- Sober, Elliot. *Evidence and Evolution: The Logic Behind the Science*. Cambridge: Cambridge University Press, 2008.
- Tichý, Pavel. "On Popper's Definition of Verisimilitude." *The British Journal for the Philosophy of Science* 25, No. 2 (1974): 155-160, <http://www.jstor.org/stable/686819>.
- Van Fraassen, Bas. *The Scientific Image*. Oxford: Oxford University Press, 1980.

SOLVING FREGE'S SUBSTITUTION PUZZLE: ANALYZING IT IN LIGHT OF DESCRIPTIVISM AND DIRECT REFERENCE THEORY

Ayesha Rehman, City College of New York

Abstract

Although replacing one proper name with another that refers to the same person does not change the truth-value of a declarative statement, it affects the truth-value of propositional attitude reports, which are cognitive relations that people hold towards propositions. Frege's Substitution Puzzle about propositional attitude reports essentially asks an important question: if two proper names co-refer in a certain linguistic community, then why does their intersubstitutability produce propositional attitude reports (that contain those proper names) with opposite truth-values? This paper attempts to explain how Description Theory of Names and Direct Reference Theory, two theories of proper names, solve Frege's Substitution Puzzle. According to the Description Theory of Names, a proper name has both a sense and a reference. In other words, a proper name expresses its sense as a descriptive meaning and it designates a specific referent. Descriptivists solve the puzzle by rejecting the Principle of Intersubstitutability of names due to their reference shift in attitude contexts; because two proper names do not entail the same sense, they cannot co-refer in attitude context and therefore are not intersubstitutable in indirect discourse. Contrary to the Description Theory of Names, Direct Reference Theory argues that a proper name is a rigid designator without any connotative attributes. It simply picks out objects and living things in possible worlds. Direct Reference Theorists solve the puzzle by stating that the seeming contradiction in the truth-values of propositional attitude reports containing co-referential names occurs because of the differing truth-values of the pragmatically implicated statements. Furthermore, this essay concludes with an argument for why Direct Reference Theory is a stronger view than Description Theory of Names.

Solving Frege's Substitution Puzzle: Analyzing it in Light of Descriptivism and Direct Reference Theory

This paper examines an example of Gottlob Frege's Substitution Puzzle about propositional attitude reports in view of two theories of proper names. Both the Description

Theory of Names and Direct Reference Theory give us an explanation of the semantic value of a proper name. While Descriptivists such as Frege and John Searle assert that names have both a sense and a referent,⁴² Direct Reference Theorists such as Saul Kripke and John Stuart Mill state that proper names simply pick out individuals/objects in possible worlds.⁴³ First, I will explain Frege's Substitution Puzzle about belief reports using the Superman/Clark Kent example. Second, I will lay out both the Description Theory of Names and Direct Reference Theory, and will explain how each theory solves the puzzle. Third, I will argue that Direct Reference Theory is a stronger view than Descriptivism.

Frege's Substitution Puzzle about Belief Reports

In the fictional American city of Metropolis, Superman and Clark Kent are the same person. They have the same reference, which is such that the identity relation 'Superman = Clark Kent' holds. If Superman and Clark Kent are alternative names for the same individual in Metropolis, then replacing one proper name for the other should be *salva veritate*. This replacement should not affect the truth-value of a declarative statement that is relevant to the linguistic environment or context of Metropolis.⁴⁴ The principle that expresses this can be formulated as follows:

Principle of Intersubstitutability: If a and b are co-referential proper names in a language L , then any true statement S of L , that contains a , can be turned into a true statement S_1 of L by replacing a with b , and similarly, any true statement S_2 of L , that contains b , can be converted into a true statement S_3 of L by replacing b with a .⁴⁵ Consider the following sentences:

- (1) Superman can fly.
- (2) Clark Kent can fly.

If one can, according to the Principle of Intersubstitutability, replace 'Superman' with 'Clark Kent' in (1), then (2) should have the same truth-value. However, this is not always the case. In cases of indirect discourse, where phrases such as 'I said,' 'John believes,' 'Emma knows,' 'they imagine' etc. are employed in the beginning, replacing co-referential expressions yield contradictory propositions. Consider:

- (3) Lois Lane believes that Superman can fly.
- (4) Lois Lane believes that Clark Kent can fly.

In the world of Metropolis, (3) is true and (4) is false on an intuitive level. Frege noticed that in propositional attitude reports such as the ones given above, substitution of co-referential names does not result in sentences with the same truth-values.⁴⁶ In "Attitude Reports: Do You Mind the Gap?" Berit Brogaard defines propositional attitudes or attitude reports as reports

⁴² Sam Cumming, "Names," *Stanford Encyclopedia of Philosophy* (2008): 3.

⁴³ Cumming 2.

⁴⁴ Berit Brogaard, "Attitude Reports: Do You Mind the Gap?" *Philosophy Compass* (2008): 93.

⁴⁵ Max Deutsch, "The Paderewski Puzzle And the Principle of Substitution," *Grazer Philosophische Studien* (2011): 123.

⁴⁶ Edward N. Zalta, "Gottlob Frege," *Stanford Encyclopedia of Philosophy* (1995): 18.

about peoples' states of mind.⁴⁷ In other words, propositional attitudes are cognitive relations that people hold towards propositions, which are truth-evaluable statements.⁴⁸ Propositional attitude verbs such as 'believe', 'know', 'think', 'fear', 'like' etc. are uttered before propositions or *that*-clauses. For instance, in sentence (3) or (4), Lois Lane's cognitive relation to the proposition 'Superman can fly' or 'Clark Kent can fly' respectively can be expressed by the propositional attitude verb *believes*. Frege's Substitution Puzzle therefore poses this important question: if two proper names co-refer in a certain environment, then why does their intersubstitutability produce propositional attitude reports (that contain those proper names) with opposite truth-values? The solution to this puzzle is important because it can help one understand the connection between thoughts and mental states, and language.

Descriptivism

In the article titled, "On Sense and Reference," Frege explains that a proper name (i.e. word, sign, sign combination, expression) such as 'the morning star' or 'the evening star' has both a sense and a reference. The proper name *expresses* its sense as a descriptive meaning, which is public and therefore can be understood by two or more people in the same way.⁴⁹ Furthermore, the name *designates* a specific referent, which is also not personal.⁵⁰ In the above example, the names 'Superman' and 'Clark Kent' have different senses in that both expressions give in a different mode of presentation, but they both refer to the same person.⁵¹ In other words, the cognitive significance that the names 'Superman' and 'Clark Kent' evoke is different. Frege asserts that every meaningful expression has a sense, but it is not necessary that it also have a reference. For instance, the sign 'Superman' has a sense, but no reference in the real world, considering fictional characters are not real in this world. Frege's theory on sense and reference is the basis for The Description Theory of Names. Also known as Descriptivism, this theory states that the semantic value of a name is some definite description 'the F'.⁵² For example, the name 'Superman' might have a semantic value of 'the superhero who can fly'. To fix the problem of not having a semantic value for people with no famous deeds or inanimate objects or imaginary beings, Descriptivism allows for a disjunction of a group of predicates; this is called Cluster Descriptivism.⁵³ As emphasized before, it is a two-element view, which asserts that names have both sense and referent. The meaning is a cluster of descriptions associated with the name while the referent is the object/living thing that satisfies all or most of the descriptions. In Kripke's excerpts from *Naming and Necessity*, Descriptivism is summed up in the following six theses:

⁴⁷ Brogaard 93.

⁴⁸ Thomas McKay, and Michael Nelson, "Propositional Attitude Reports," *The Stanford Encyclopedia of Philosophy* (2000): 1.

⁴⁹ Gottlob Frege, "On Sense and Reference," *The Philosophy of Language*, Eds. A.P. Martinich and David Sosa (2013): 36.

⁵⁰ Frege 37.

⁵¹ Frege 35.

⁵² Cumming 4.

⁵³ Cumming 5.

- i. To every name 'X', there corresponds a cluster of properties, the family of those properties F such that a speaker A believes 'FX'.
- ii. One of the properties, or some conjointly, are believed by A to pick out some individual uniquely.
- iii. If most, or a weighted most, of the F's are satisfied by a unique object y, then y is the referent of 'X'.
- iv. If the vote yields no unique object, 'X' does not refer.
- v. The 'If X exists, then X has most of the F's' is known a priori to A.'
- vi. The statement that 'If X exists, then X has most of the F's' expresses a necessary truth.⁵⁴

To solve the Substitution Puzzle, Frege argues that in indirect discourse, 'Superman can fly' and 'Clark Kent can fly' refers to its customary sense (thought) rather than its reference (a truth-value). The sense of a name, which is fine-grained and is therefore able to convey more knowledge than the truth-value alone, is what determines its referent. In fact, if two names have the same sense, then they have the same referent. However, it is not necessary for a referent to have the same sense. The difference in the sense of 'Superman' and 'Clark Kent' explains the difference in the truth-values of (3) and (4), where the sense or way of presentation of 'Superman' is a superhero who flies and the sense of 'Clark Kent' is a bespectacled reporter for the "Daily Planet". Because the concepts of 'Superman' and 'Clark Kent' do not entail the same sense, they cannot co-refer in attitude context and therefore are not intersubstitutable in indirect discourse. In other words, in propositional attitude reports, one has to reject the Principle of Intersubstitutability of names due to their reference shift. This is how Frege solves the puzzle.

Direct Reference Theory, Millianism & Neo-Russellianism

Direct Reference Theory or Millianism proposes that a proper name has a referent only. Strictly speaking, it is a rigid designator. This means that a proper name picks out the same object or person in all possible worlds where that object or person exists. Direct Reference Theorists such as Kripke posit that the reference is established through a dubbing or creation event, where a name is given and which spreads by a causal chain of reference. Kripke elaborates the Casual Theory:

An initial 'baptism' takes place. Here the object may be named by ostension, or the reference of the name may be fixed by a description. When the name is 'passed from link to link', the receiver of the name must, I think, intend when he learns it to use it in the same reference as the man from whom he heard it. If I hear the name 'Napoleon' and decide it would be a nice name for my pet aardvark, I do not satisfy this condition.⁵⁵

⁵⁴ Saul Kripke, *Naming and Necessity*, (Cambridge, Mass.: Harvard University Press, 1980) 53.

⁵⁵ Kripke, *Naming and Necessity*, 63.

A speaker, using a name "NN" on a particular occasion is denoting or referring to some item 'x' if there is a causal chain of reference preserving links leading back from the speaker's use ultimately to the item 'x' itself being involved in a name-acquiring "baptism."

In the article titled, "Of Names," Mill defines proper names in a similar fashion. He states that proper names are not connotative, that is, they do not imply any attribute. Proper names only signify a specific subject. Mill writes, "Proper names are attached to objects themselves and are not dependent on the continuance of any attribute of the object...."⁵⁶ One can think of names as tags that do not provide any additional information. Mill goes on to say, "It may be said, indeed, that we must have had some reason for giving them those names rather than any others, and this is true, but the name, once again, is independent of the reason." Even if the dubbing event has a reason behind it, the name given to that individual does not carry attributes attached to that individual. This is consistent with Kripke's example of naming a pet after someone famous like Napoleon. If I name my pet cat 'Einstein,' because it behaves in a very clever way, it is not rational for me to begin connoting definite descriptions about superior intelligence with my cat's name. That would be silly!

Now that Direct Reference Theory has been laid out, one can see that both (3) and (4) are true. However, how does this theory explain that Lois Lane believes that Clark Kent can fly, even if she does not believe the sentence 'Clark Kent can fly'? Direct Reference Theory clarifies this by incorporating another area of the study of language: pragmatics. According to Martinich and Sosa, pragmatics is the study of what speakers do with language, that is, how speakers can perform actions with words and get across more than the words' literal meanings.⁵⁷ To solve the puzzle in the Superman/Clark Kent example, neo-Russellians make use of conversational implicatures. They explain that (3) pragmatically implicates a true statement that Lois Lane believes Superman can fly as a superhero whereas (4) pragmatically implicates a false statement that Lois Lane believes Superman can fly as a reporter.^{58, 59} Confounding pragmatics with semantics does not make (4) false. In fact, the implicature it generates is false. This is how Direct Reference Theorist or neo-Russellians explain the apparent contradiction due to intersubstitutability in attitude context.

Reasons for Upholding Direct Reference Theory

Note that, unlike Descriptivism, Direct Reference Theory does not get rid of the Principle of Intersubstitutability for solving Frege's puzzle, but rather makes use of the truth-value of implicatures to explain why a rational agent might appear to both assent and not assent to the same proposition simultaneously. Moreover, the Descriptivists' reason for rejecting the Principle of Intersubstitutability and thus solving Frege's Substitution Puzzle can be shown to be unconvincing when one looks at Kripke's Paderewski Puzzle. This is because the Paderewski Puzzle, which also involves propositional attitude reports, is not a Substitution Puzzle and therefore cannot be solved by denying Intersubstitutability.^{60, 61} Consider:

⁵⁶ John Stuart Mill, "Of Names," *The Philosophy of Language*, Eds. A.P. Martinich and David Sosa (2013): 32.

⁵⁷ A.P. Martinich, and David Sosa, "Introduction," *The Philosophy of Language* (2013): 2.

⁵⁸ McKay 9.

⁵⁹ Brogaard 97.

⁶⁰ Saul Kripke, "A Puzzle About Beliefs," *Meaning and Use* (1979): 449.

⁶¹ Brogaard 97.

- 5) Peter believes that Paderewski has musical talent.
- 6) Peter disbelieves that Paderewski has musical talent.

Suppose Peter comes to know 'Paderewski' as the famous Polish pianist, so obviously, he assents to the statement 'Paderewski has musical talent.' In another context, Peter learns of 'Paderewski' who was the Polish Nationalist leader and prime minister, so after identifying 'Paderewski' as the Polish politician, Peter assents to the sentence 'Paderewski has no musical talent'. According to Kripke, (6) and (7) may be both true under different circumstances because Peter fails to realize that Paderewski, the pianist, is the same person as Paderewski, the politician (who Peter believes cannot have musical talents by virtue of his statesmanship).⁶² Kripke notes that rejecting Intersubstitutability would not solve this puzzle because there is no replacement of co-referential terms! Instead, some other way has to be devised. Because Descriptivists deny Intersubstitutability in propositional attitude reports, Kripke asserts that they must present a serious argument as to why Intersubstitutability has to be rejected. This is not to imply that a solution cannot reject Intersubstitutability, but rather that this move would not be helpful for Kripke's Paderewski Puzzle.⁶³

Besides providing a better solution to the Substitution Puzzle, I also think that Direct Reference Theory provides a better view of the semantic value of proper names than Descriptivism. According to Kripke's modal argument, proper names are rigid designators while definite descriptions are not. If that is the case, names cannot have the same meaning as definite description(s) associated with it.⁶⁴

In addition to that, one might not know any descriptions associated with a proper name. However, that does not mean that proper names have no meaning. Even if one can identify descriptions for a proper name, one cannot be certain which description picks out a unique object/person. For instance, the description for 'Aristotle' could be arbitrary descriptions such as 'the author of *Nicomachean Ethics*' or 'the greatest student of Plato' etc. There is no way to identify a single description or a cluster that actually assigns meaning to 'Aristotle'. Furthermore, people often hold wrong descriptions for a proper name. For example, some people believe that 'Einstein' is 'the creator of atomic bomb.' Despite the fact that this definite description is incorrect, people are pinpointing a unique individual in history. Yet, according to Descriptivism, the referent 'the creator of atomic bomb' should be Oppenheimer.

Although Direct Reference Theory blurs the demarcation line between semantics and pragmatics by bringing up implicatures, one has to realize that that distinction has been difficult to pinpoint. The principal semantic notions are truth and reference, but including an analysis of pragmatics provides a full picture.⁶⁵

Descriptivism seems to make the leap that the Fregean sense transfers information upon its utterance without a clear justification. Direct Reference theorists recognize the fact that the dubbing event of a proper name might be motivated by a reason, but are prudent not

⁶² Kripke, "A Puzzle About Beliefs," 449.

⁶³ Brogaard 97.

⁶⁴ Cumming 6.

⁶⁵ Martinich 2.

to conclude that some sort of knowledge is embedded within the name itself as a result of the dubbing.

It seems that Kripke is applying David Lewis' view of convention of truthfulness and trust for the Casual Theory. In Lewis' view, being truthful roughly translates to saying things that one thinks are true and being trusting means that one believes others' utterances to be true. According to Lewis, the following six conditions must be satisfied for a convention of truthfulness/trust in L to prevail:

- i. Everyone conforms to a convention of truthfulness/trust in L.
- ii. Everyone believes that the others conform to truthfulness/trust in L.
- iii. The belief that others conform to truthfulness/trust in L gives everyone a good and decisive reason to conform to truthfulness/trust in L themselves.
- iv. There is a preference for general conformity to truthfulness/trust in L rather than slightly-less-than general conformity to truthfulness/trust in L.
- v. There is at least one alternative regularity, truthfulness/trust in L', such that condition 3 and 4 hold for L', and such that there is no way to conform to truthfulness/trust in L and L' at the same time.
- vi. Conditions 1-5 are common knowledge in a population P.⁶⁶

Assuming that Kripke agrees with this definition of coordination convention being used in the chain of reference, then it adequately addresses reference shift examples (e.g. about Madagascar once being known as a portion of mainland Africa, but then undergoing a reference shift after Marco Polo took it to refer to the great African island), brought up by Gareth Evans in "The Casual Theory of Names".⁶⁷ This is because the conformity in use of a proper name for a unique person/object would not allow for reference shift. Suppose Kripke's Casual Theory does not depend on coordination convention, then his insistence on counting speaker intention and audience recognition and execution of that intention while using proper names will be enough to counter reference shift examples. In view of the above reasons, Direct Reference Theory overrides Descriptivism.

⁶⁶ David Lewis, "Languages and Language," *The Philosophy of Language*, Eds. A.P. Martinich and David Sosa (2013): 684-686.

⁶⁷ Gareth Evans, "The Casual Theory of Names," *The Philosophy of Language*, Eds. A.P. Martinich and David Sosa (2013): 79.

References

- Brogaard, Berit. "Attitude Reports: Do You Mind the Gap?" *Philosophy Compass*, 3.1 (2008): 93-118. Web. 15 Dec. 2014.
- Cumming, Sam. "Names." *Stanford Encyclopedia of Philosophy*. Stanford University, 17 Sept. 2008. Web. 15 Dec. 2014.
- Deutsch, Max. "The Paderewski Puzzle And the Principle of Substitution." *Grazer Philosophische Studien*. 83 (2011): 123-141. Web. 15 Dec. 2014.
- Evans, Gareth. "The Casual Theory of Names." *The Philosophy of Language*. Eds. A.P. Martinich and David Sosa. New York: Oxford University Press, 2013. 74-85. Print.
- Frege, Gottlob. "On Sense and Reference." *The Philosophy of Language*. Eds. A.P. Martinich and David Sosa. New York: Oxford University Press, 2013. 35-47. Print.
- Lewis, David. "Languages and Language." *The Philosophy of Language*. Eds. A.P. Martinich and David Sosa. New York: Oxford University Press, 2013. 682-700. Print.
- Kripke, Saul. "A Puzzle About Beliefs." *Meaning and Use*. Ed. A. Margalet. Dordrecht: D. Reidel Publishing, 1979. 433-459. Print.
- Kripke, Saul. From *Naming and Necessity*. Cambridge, Mass.: Harvard University Press, 1980. 53-63. Print.
- Martinich, A.P. and Sosa, David. Introduction. *The Philosophy of Language*. By Martinich. New York: Oxford University Press, 2013. 1-25. Print.
- McKay, Thomas & Nelson, Michael. "Propositional Attitude Reports." *The Stanford Encyclopedia of Philosophy*. Stanford University, 16 Feb. 2000. Web. 9 Dec. 2014.
- Mill, John Stuart. "Of Names." *The Philosophy of Language*. Eds. A.P. Martinich and David Sosa. New York: Oxford University Press, 2013. 29-34. Print.
- Zalta, Edward N. "Gottlob Frege." *Stanford Encyclopedia of Philosophy*. Stanford University, 14 Sept. 1995. Web. 14 Dec. 2014.

SEARCHING FOR ETHICS' GROUNDING: A CASE FOR MORAL FEELING AND THE HUMAN RELATIONSHIP TO NATURE

Katie Coulter, Eastern Michigan University

Abstract

The following essay considers the question of how ethical and moral theories are possible in conjunction with the "death of God" as conceptualized by Nietzsche and other continental thinkers. I argue that ethical and moral action become possible through, and require, a deep affective experience of something as having absolute value, and that this kind of experience of absolute value can be found in human beings' relationship with nature. Using the work of Bernard Williams and John Russon, I argue that the climate crisis facing the planet makes apparent this relationship, and makes possible a particular kind of affective response to nature which, in turn, makes ethical action possible.

Searching for Ethics' Grounding: A Case for Moral Feeling and the Human Relationship to Nature

As Nietzsche heralded the death of God, he identified a number of consequences of this intellectual event. First, Nietzsche celebrated the end of the idea that ethics and morality are determined and handed down by a deity, as well as the sweeping aside of the idea that in order for one to be good, there must be a moral authority as the source of what is good. With academic scholarship and scientific investigation dismantling and replacing foundational aspects of Christianity and religious belief, Nietzsche saw "the collapse of any theistic support for morality" (Crowell), and that "the belief in the Christian God has become unworthy of belief" (Nietzsche, 67). For Nietzsche, the end of the notion of a divine source of morality and absolute value was "a liberating opportunity to take responsibility for meaning, to exercise creativity" (Crowell). Without belief in a divine power determining morality, people are free and responsible to formulate their own conception of moral action and their own attribution of value. Nietzsche describes this freedom thus:

. . . we philosophers . . . feel, when we hear the news that “the old god is dead,” as if a new dawn shone on us . . . At long last the horizon appears free to us again, even if it should not be as bright; at long last our ships may venture out again . . . all the daring of the lover of knowledge is permitted again; the sea, *our* sea lies open again, perhaps there as never yet been such an “open sea” (68).

With the death of God, ethics is placed firmly where it should be, and its source is acknowledged to be what it is and always has been according to Nietzsche: within human beings. For Nietzsche, there is nothing behind value judgments other than one's own will (Leiter). While in the end the view is more complex than this, the important take away for this paper is that the death of God is the death of the idea that there is objective or absolute value. Nietzsche had his own ideas about what ethics and morality should look like in the face of this, but for the purpose of my argument the death of God presents both a loss and gain: the loss of a millennia-old source of absolute value, and the gain of one's agency (and the recognition of that agency) to determine for oneself what is good and bad, right and wrong.

With the loss of objective value comes the prospect of nihilism and the pain and confusion that can result from it. Several decades before Nietzsche, Hegel described this pain: “The pure concept, however, or infinity, as the abyss into which all being sinks, must characterize the infinite pain . . . the feeling that God Himself is dead” (Groom, Fritz, 29). Nietzsche identifies early on in his writing a “shadow”, as though “some sun seems to have set and some ancient and profound trust has been turned into doubt” (68). Rose Pfeffer provides a good understand of the predicament facing human beings:

With the loss of a sense of purpose, resulting from the denial of a teleological universe, the foundation of a moral world order is shattered. Man (*sic*) no longer possesses the ideals and absolute goals toward which to strive. He (*sic*) has lost all direction and purpose . . . He (*sic*) is lost, without a God and without the promise of a better world.

Having lost the most readily available source of absolute value, one falls into an infinity of possible values with no handhold. The feeling of responsibility that comes with the freedom to determine value, and thus morality, for oneself, can be paralyzing. Each thing encountered or considered must be evaluated independently and its value sought by each individual moral agent for themselves. With no solid prescriptions of value, Hegel's abyss opens, and the prospect that nothing has value looms. I contend however, that additionally distressing is the endless internal search for something on which to base one's system of value, meaning, and morality. The turmoil and confusion of this search for moral solidity has no equal, for without moral solidity, coherence and meaning cannot be built and relief from the search cannot be found.

At this point in the discussion, I find it necessary to make apparent an important distinction. I am not arguing that in order to achieve an ultimate grounding for ethics and morality there must in fact be something of absolute value. I do not mean to suggest that morality necessarily requires an objective value. On the contrary, I believe, as Nietzsche did, that behind morality there is nothing but our own human formulations of value, and behind

these formulations of value is moral feeling. I am arguing that the *experience* of something as having an absolute value, whether positive or negative value, is essential to constructing a meaningful understanding of the world, and therefore for constructing an ethical theory for moral action. At its foundation, ethics is a way to find the relative importance of various things under consideration – an ethical dilemma is the struggle to determine what, in a given situation, is most important.

Back to the problem at hand: with the death of God, we have found ourselves without our most familiar source of absolute value, and as the above distinction clarifies, what we have truly lost is the most readily available source of the *experience* of absolute value. There are plenty of ethical theories offering their best understanding of what is most important and how that can be determined. Various deontologies, utility principles and virtue systems offer accounts of what is the most important good, and yet often the question of how they are ultimately grounded remains unanswered. To illustrate this, it is worth looking at how the Euthyphro Dilemma has been extended to any systematic ethics. The Euthyphro Dilemma finds religion's account of morality to be without substance because it either 1) determines the good, in which case it could decree things normatively considered to be immoral, such as murder, to be moral, or it 2) merely identifies the good, in which case something else more foundational must function as the grounding for determining what is moral. Mark Taylor summarizes the point thus:

Systematic ethics, by their nature, identify almost all moral obligations as contingencies that rely on an ultimate self-sufficient principle. Such a principle is reputedly good by its nature and serves as the anchor point from which all other duties originate. In fact, the rest of the system is really just an extended explication of the foundational principle. If we were to find that the anchor point is *not* independent or necessary, then we should reject that whole system (46).

The problem, Taylor contends, is that all systems of ethics fall victim to the Euthyphro Dilemma – in the case of consequentialism, Taylor concludes (after much argumentation that will not be covered here) that “(UP)[the utility principle] is the foundation of Utilitarian morality, and there exist counter-examples showing that (UP) cannot be equivalent to moral goodness, so (UP) and Utilitarianism are not related ontologically to moral goodness” (50). All of this is to say that the problem of experiencing Hegel's abyss or Nietzsche's nihilism at the loss of the experience of absolute value is not easily solved by other groundings for ethics, and that if this phenomenological experience of chaos and an infinity of moral ambiguity cannot be given some kind of handhold or foundation, then substantive moral action becomes at best exceedingly difficult and at worst inconceivable. There is a phenomenological element to ethics that, as with many phenomenological insights, goes overlooked and yet must always already be the case in order for moral action to occur at an individual level: one must *feel* that something is important in order for one to be moved to act. To be intellectually convinced is to respond to a strong argument, but more basically, to be convinced is a *feeling* and an *experience*. One can think that something is important, but unless one also experiences it as important, then impetus for action will be extraordinarily difficult to come by. The conclusion is this: without a

moral feeling to arouse, galvanize, and thus provoke action, the best ethical system (whether deontology, consequentialism, etc.) is impotent.

This is illustrated by considering a mental illness like depression. A person with severe enough depression will find motivation for actions of any kind significantly difficult. Although cognitive capacities can be affected by depression, what is most handicapped is one's affective responses. The world is not experienced as significant, important, or meaningful – without the feeling that things have importance, the depressed person often does not feel any incentive to carry out a project of any kind. The result can be that the person's rational capacities are entirely unaffected, but even the most carefully constructed argument for actions of any kind, let alone moral ones, are not convincing to the point of catalyzing action. The depressed person does not feel or experience the importance of a thing, and thus is unable to generate sufficient motivation. Rational thought and strong argument alone are an insufficient grounding for an ethical system, because rational arguments do not fulfill the requirement of feeling morally moved. Good arguments can contribute to or cause one's affective response – a good argument can be the thing that makes one experience the value of a thing. But the catalyst for action remains the *experience* of value.

I find an interesting source of support in Robert Elliot's book *Faking Nature*. Elliot puts forth a very complex and careful metaethical theory and grounding for value. Elliot is exceedingly careful to avoid doing exactly what I am proposing – Elliot wants his argument to be solidly grounded on a principle that is completely self-sufficient, and he seeks to justify in this manner all his claims of natural value. And yet, his whole account of value essentially rests on one footnote: "That nature has value is, so to speak, a brute value fact. Although the fact does not admit of further explanation, it requires emphasis and discussion . . ." (Elliot, 157). What Elliot is asserting is contrary to his intended project of finding independent and necessary value that can avoid the pitfalls of the Euthyphro Dilemma. A "brute value fact" is nothing if not something that "just has" value. The claim that something "just has" value is an affective claim. It is feeling and experiencing some thing as important and valuable. In short, it is experiencing absolute value and thus a handhold while falling into Hegel's abyss.

While the experience of something as having value is dismissed as a foundation for ethics because it is capricious, lacking rigor, or far too relative, I would like to contend that the experience of something having value is in fact one of the best possible groundings for ethics, and as discussed above, possibly a requirement for engaging in moral action. The criticisms of capriciousness, lack of rigor and relativity are serious and require discussion, however. A further elucidation of what it means to experience something as having absolute value will help to dispel these worries.

To experience something as having absolute value, there can be no ambiguity at all in that particular experience. Absolute value, or in Elliot's words a "brute value fact", implies an all-encompassing certainty about the value relationship between oneself and the thing experienced. It is not the case that anything we value satisfies this feeling of encompassing certainty. If deeply and thoroughly considered, nearly anything experienced as having value can admit of significant ambiguity – even the value of those one loves most can be consumed and questioned in the yawning jaws of nihilism. This fact is precisely why nihilism is so persistent: What is left is a world of mere appearance and semblance, possessing no certainty or permanence, having no goals, no unity, no truth, no being. "A "horror vacui" seizes man (*sic*) . . .

those “higher values” which the Platonic-Christian tradition falsely endowed with objective validity . . . are in fact merely subjective categories” (Pfeffer 76, 77).

If carefully constructed analytical ethical theories are not enough to convince one of something's value to the point of inspiring substantive action and a handhold in the abyss, then what would be enough to do this while also avoiding being “falsely endowed with objective validity” (Pfeffer)? We can find this very thing in the human relationship with nature. I would like to propose that our relationship to nature, while superficially ambiguous, in is fact far more essential than we generally take it to be, and that the climate crisis makes this essential, given relationship apparent again. Within our complex, technological and domination-based experience of nature, there is a more fundamental, foundational relationship that, though obscured, is in fact original. Through discussion of the work of Bernard Williams and John Russon, I will offer my case for this original givenness as that to which we can turn for the experience of absolute value and a handhold as the abyss of nihilism opens beneath us and our human search for some absolute value troubles us ever more.

We are searching for an experience so powerful, complete, and unambiguous that it serves as a source of the experience of absolute value, and therefore as something solid on which we can build our understanding of the value and importance of other things in our experience. Our relationship with nature satisfies this kind of complete, unambiguous relationship, as John Russon describes:

There is the inexplicable nurturance of the sun . . . and of the earth that is the foundation of stability and consistency – these are two original senses, irreducible forms, that appear compellingly and guidingly for us. These are forms to which we are inexplicably attuned and *to which we owe everything* [emphasis added]. This can be said for the world of nature . . . in general. We only ever occur ourselves *within* the self-occurring realm of nature . . . The fertile earth, the sky that supplies nurturing warmth and clarifying light, and the self-sufficing rhythm of growth, death, and regeneration are not senses we invent or realities we make. *It is only within their context that we occur* [emphasis added] (23).

Here Russon highlights our givenness as biological beings on a natural planet. We fit within the biological processes and natural realities of earth, sky, growth and death in a way that is simply unavoidable; these processes cannot be circumvented. The necessary conditions for our very existence are not contingent – life has developed on the planet in a particular way, and as such it requires and owes everything it is to the given reality of the natural world. We have certain kinds of bodies – mortal bodies, bodies originating in nature. This givenness is not ambiguous, it is not partial, it cannot be questioned. It cannot be questioned because it is the very parameters by which we exist at all, and “It is these realities to which we must answer, and their very reality entails that we will be ruined if we fail to respect them” (Russon, 23). Our relationship to nature is originary in the sense that it is on the basis of nature that we even have the power to question our relationship with nature at all.

Our reality [is] that which exercises its wonderful (. . . both great and terrible) power always within a context of other given wonders, to which we must bring

the appropriate level of honor and respect. All of our accomplishments occur within and in the terms of this *given* world, through our *given* powers (especially our “cunning” power to control nature by turning its powers against itself). That we are definitively constrained by the givenness is shown by the ineffaceability of our death. (Russon, 25)

There is a very specific, unambiguous way in which we are related to nature: nature is that by which there is anything at all, including us. And as Russon indicates, if we do question or doubt this relationship, we do so at our own peril. We may question our relationship to nature and act with hubris in ways that flout what can only be described as a holy bond, but doing so will only bring us more firmly to the absolute value, and to perhaps the only thing that cannot be circumvented.

Bernard Williams gives hints to this kind of relationship in his essay *Must a Concern for the Environment be Centred on Human Beings?* Though Williams’ objective is different than my own, at the end of his argument he considers the idea that “human beings have two basic kinds of emotional relations to nature: gratitude and a sense of peace, on the one hand, terror and stimulation on the other” (238). He goes on to talk about “what might be called *Promethean fear*, a fear of taking too lightly or inconsiderately our relations to nature . . . a sense of an opposition between ourselves and nature, as an old, unbounded and potentially dangerous enemy, which requires respect” (239). He then identifies what he considers to be important about this affective response to nature:

We should not think that if the basis of our sentiments is of such a kind, then it is simply an archaic remnant which we can ignore. For, first, Promethean fear is a good general warning device, reminding us still appropriately of what we may properly fear. But apart from that if it is something that many people deeply feel, then it is something that is likely to be pervasively connected to things that we value, to what gives life the kinds of significance that it has. (239)

This Promethean fear that Williams describes is just the kind of affective response that admits of no ambiguity. The deep-seated, pervasive wariness and respect for nature as both our genesis and the source of our mortality is not the kind of thing that can fall victim to nihilism. Kant illustrates this power of nature in his consideration of the sublime:

Bold, overhanging, as it were threatening cliffs, thunder clouds towering up into the heavens, bringing with them flashes of lightning and crashes of thunder, volcanoes with their all-destroying violence, hurricanes with the devastation they leave behind, the boundless ocean set into a rage, a lofty waterfall on a mighty river, etc., make our capacity to resist into an insignificant trifle in comparison with their power. (144)

If one does question or disregard the value of nature, what nature *is to us* or *means to us*, one will quickly and surely feel the bite of that mistake: you cannot disregard your biological need for water and food or you will not survive, you cannot disregard the power of the oceans

or you will drown, you cannot disregard the force of the wind or you will be battered, you cannot find the freezing temperature of winter “mere appearance and semblance” (Pfeffer, 76) or you will freeze. One cannot be complacent in the face of the absolute givenness of nature – there will be swift consequences to equivocating about the value relationship between oneself and nature. The abyss of not knowing *what something means* to you or for you suddenly has a hundred handholds in the form of things that have specific and an absolute value, whether positive or negative, to your survival as a general human being, and also to the survival of the specific human body that is you.

And yet there remains an important question: is this the experience of nature that humanity has now? It seems undeniable that our givenness appears more and more as relativity, contingency. Technology increasingly pervades our life, domination of nature has only become more prevalent – at the most extreme in the United States, many of us live constantly in climate controlled dwellings, never experiencing extreme heat or cold for long. We pipe water into the desert and grow manicured green lawns. We have available all manner of food at all times of year. Everything can seem to be possibly contingent; anything can be circumvented if one employs enough cunning. Indeed, even our ties to the planet itself seem to be arbitrary – there have been human beings continuously living off of the planet on the International Space Station in the void of space for over fifteen years. If we assert our human powers enough, it seems as though we are subject to nothing, answerable to nothing, falling once more in a an abyss of an infinity of possible meanings; once more nothing is absolute. But these examples betray themselves. There is nowhere where our utterly unconditional need for oxygen and atmospheric pressure are more urgently palpable than when venturing into space. And back on earth, though for long we have evaded, questioned, and circumvented what Russon calls that “power to which we owe whatever we are,” our answerability, our ultimate givenness is moving back into our awareness in the form of the climate crisis. We cannot control the rising seas, the migrating climates, the droughts, the increasingly vicious storms, the toxic air, all of which we have some amount of responsibility for. Nature is reasserting itself in our experience as Williams’ “old, unbounded and potentially dangerous enemy,” one which requires our respect, lest we risk our own ruin.

Pushing the limits only brings into sharper focus that to which we are truly subject, that which is not contingent, not ambiguous, and cannot be circumvented. That which is once again experienced as absolute – our given relationship to nature as both our origination and our potential destruction – can be our handhold, it can be a grounding to meaning, the source of experiences and feelings of significance, and therefore the impetus for moral action that we are searching for.

References

- Crowell, Steven, "Existentialism", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.). Web. 16 December 2015.
- Elliot, Robert. *Faking Nature*. New York: Routledge, 1997. Print.
- Groom, S. Alyssa, and Janie M. Harden Fritz. *Communication Ethics and Crisis: Negotiating Differences in Public and Private Spheres*. Madison: Fairleigh Dickinson UP, 2012. Print.
- Leiter, Brian, "Nietzsche's Moral and Political Philosophy", *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition), Edward N. Zalta (ed.). Web. 16 December 2015.
- Nietzsche, Friedrich. *Existentialism*. Ed. Robert C. Solomon. New York: Oxford University Press, 2005. 65-101. Print.
- Pfeffer, Rose. *Nietzsche: Disciple of Dionysus*. Cranbury: Associated University Presses, Inc., 1972. Print.
- Russon, John. *Bearing Witness to Epiphany*. Albany: State University of New York Press, 2009. Print.
- Taylor, Mark. "The Euthyphro Dilemma and Utilitarianism." *Acta Cogitata*. 3 (2015) : 43-51. Web. 15 December 2015.
- Williams, Bernard. "Must a concern for the environment be centred on human beings?", *Making Sense of Humanity*. 1st ed. Cambridge: Cambridge University Press, 1995. pp. 233-240. *Cambridge Books Online*. Web. 16 December 2015.