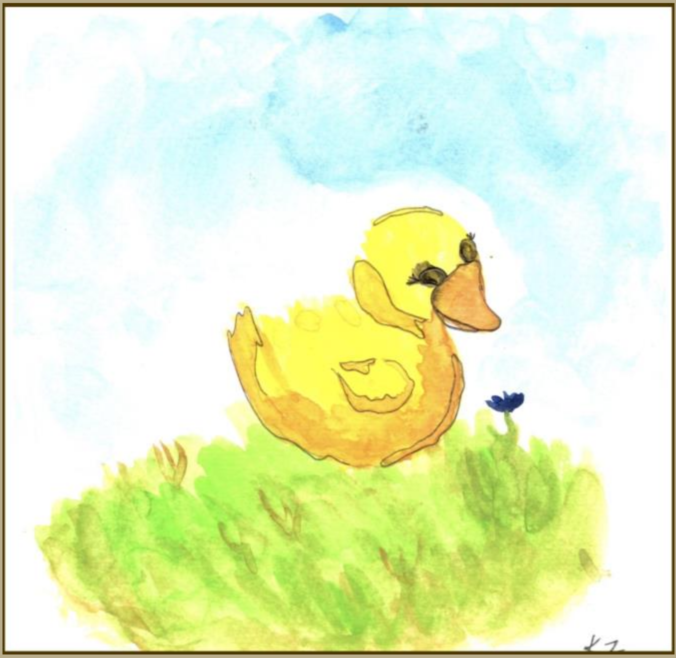
Philosophy in Practice

VOLUME 20 – SPRING 2026



CALIFORNIA STATE UNIVERSITY, LOS ANGELES
DEPARTMENT OF PHILOSOPHY

Philosophy in Practice

VOLUME 20 - SPRING 2026

CONTENTS

ACKNOWLEDGEMENTS

Philosophy in Practice is produced by students at California State University, Los Angeles. The editorial staff wishes to thank all the authors for submitting and refereeing articles, the faculty of the philosophy department for agreeing to supervise individual contributions, and the University for its generous financial support.

California State University, Los Angeles
Philosophy Faculty

David Pitt (2003-), Chair, Ph.D. City University of New York. Philosophy of Mind, Philosophy of Language, Metaphysics

Mohammed Abed (2008-), Ph.D. University of Wisconsin, Madison. Ethics, Social and Political Philosophy, Philosophies of Violence, Genocide and Terrorism

Bruce Atta (1996-), M.A. California State University, Los Angeles. Metaphysics, Epistemology, Ethics, Social and Political Philosophy

Mark Balaguer (1992-), Ph.D. City University of New York. Philosophy of Mathematics, Metaphysics, Meta-ethics, Philosophy of Language, Logic

Talia Bettcher (2000-), Ph.D. University of California, Los Angeles. History of Modern Philosophy, Philosophy of Self, Gender and Sexuality

Joel Chandler (2014-), M.A. California State University, Los Angeles. Virtue Ethics, Pragmatism, Existential Hermeneutics

Jay Conway (2005-), Ph.D. University of California, Riverside. 19^{th}-20^{th} Century Philosophy, Modern Philosophy, Aesthetics, Social and Political Philosophy

Richard Dean (2009-), Ph.D. University of North Carolina, Chapel Hill. Ethics, Kant's Moral Philosophy, Applied Ethics

Foad Dizadji-Bahmani (2013-), Ph.D. London School of Economics, United Kingdom. Philosophy of Science, Philosophy of Physics, Philosophy of Probability

Ronald Houts (1983-), Ph.D. University of California, Los Angeles. Metaphysics, Epistemology, Logic

Katie Howard (2023-), Ph.D. Emory University. Feminist Philosophy, Political Philosophy, Affect Theory, Decolonial Thought, History of Philosophy

Keith Kaiser (2001-), Ph.D. University of California, Los Angeles. Metaphysics, Philosophy of Mind

Sheila Price (1964-), M.A. University of California, Los Angeles. Recent Philosophy, Comparative Religions, Medical Ethics, Environmental Ethics

Michael K. Shim (2007-), Ph.D. State University of New York, Stony Brook. 20^{th} Century Continental Philosophy, Phenomenology, Husserl, Modern Philosophy, Philosophy of Mind

EMERITUS PROFESSORS

Thomas Annese (1961-1992), Epistemology, Modern Philosophy

Sharon Bishop (1967-2004), Ethics, Political Philosophy, Philosophical Psychology, Feminist Ethics

Donald Burrill (1962-1992), Ethics, Philosophy of Law, American Philosophy

Ann Garry (1969-2011), Feminist Philosophy, Philosophical Methodology, Epistemology, Applied Ethics, Wittgenstein, Philosophy of Law

Ricardo J. Gómez (1983-2011), Philosophy of Science and Technology, Philosophy of Mathematics, Kant, Latin American Philosophy

Henry R. Mendell (1983-2023), Ph.D. Stanford University. Ancient Philosophy, History of Ancient Mathematics and Science, Philosophy of Science, Metaphysics

Kayley Vernallis (1993-2019), Moral Psychology, 19th and 20th Century Continental Philosophy, Feminist Philosophy, Ethics, Aesthetics, Gender and Sexuality

George Vick (1967-1997), Metaphysics, Phenomenology, Existentialism, Philosophy of Religion, Medieval Philosophy

Collectibles and The Right Stuff

Niyaz Mahmud

Pre-Collection

I had a conversation with a friend where he told me that he bought some Beta Dual Lands while traveling.[1] If you're not familiar with Magic: The Gathering, or simply Magic,[2] it is a tabletop game played with cards. It's known as a TCG or CCG[3] which are trading card games or collectible card games respectively. In early 2026, the value for a Beta Dual Land (so called because they represent and act as two different land types in the game and never really reprinted in such an in-game-pure form) is hovering around $2000. I was surprised by the fact that my friend would spend that much money on something which is, to me, just a scrap of cardboard. I don't think my attitude is unique but I certainly don't think his is either especially if you look at the history of the game. We're not going to go into detail about the history of Magic but consider the value-hierarchy implied: My friend has a sentimental attachment to Magic, and therefore has started a collection of cards spanning decades, and thus is able to talk to people who know about Magic about his amassed collection; with whatever the social consequences of that are (probably good if you like Magic, bad if you don't). Ultimately, the fondness for Magic and social community he enters as part of that fondness. This creates a willingness to spend substantial sums of money on what are, to outsiders, pieces of flimsy cardboard with neat pictures.

So, what is really happening with this Magic collection, or any collection, and what can a collection say about who we are and what we care about? At a high level, to collect is a deliberate act of accumulating and possessing a certain thematic class of object (a material object usually, but not necessarily so), and the accumulation is the collection. Varied influences act on this desire to collect. A collection may serve as an externalization of ourselves and what we care about. 21st Century Capitalism[4] has set up structures that turn collecting from labor of passion into a labor of speculation (e.g. speculative investments that, 'Invisible-Hand' willing, go up in value). We can say that collections function as a tool to make exchanges between sentimental (or personal) value, economic (or financial or monetary) value, and cultural value (or clout). But what types of value make for a worthy collection? Plainly stated, having things is not in itself wrong, but there is a way to collect that is aboutness-intentional and deliberateness-intentional, without thoughtless or harmful consumption. The ideal collection should first spring from sentiment, necessarily. Second, if a community for the class of collection exists, have cultural value to that community. Third, it should only be concerned with economic value as a tertiary concern. For communities, the rights to access an ideal collection become important; i.e. who has access to a collection of religious artifacts, or to a collection of secrets, or especially in the case of items that have an important cultural heritage. The *simulacra* (we'll get to this soon) of collection is a shadow of the ideal, where the economic value, restricted access, or lack of sentiment turn it into something else, like pure-speculation or investment.

This paper is organized into a few sections. The first has a discussion focusing on defining a collection. This including what does and doesn't go into a collection (spoiler: almost anything you can think of can be collected in some way) and what a collection is. The second section talks about different ways we value collections. Next there is a non-exhaustive list of different kinds of collections and a bit about the function of those collections. It continues with what might point us towards a worthy collection with the last section talking about contemporary collections and collectors in a socio-economic context. Finally, I'll wrap everything up and conclude by tying it all together.

What is A Collection?

A collection is a set of objects in a class or type (usually physical objects, but not necessarily, e.g. data) that are valuable (for whatever reason). The collection must be cared for once acquired, and the use is primarily in the act of possession and display of the objects. Collecting then, is the acquisition of elements of a collection class with attention paid to what makes an element a better addition when compared to another potential element. Since it also must remain organized elements of a material collection must be culled to admit new elements if there are restraints on that collection growing. This can be less important for non-material collection but the culling aspect refines the collection and may be important when considering the time it takes to care for each element, house it, pay for its upkeep or whatever.

Obviously, collecting isn't a fringe activity; If we don't collect anything deliberately it's likely we know someone who does. Joe Ungemah, in *Fanatic*, says 30% of people self-report that they're a collector of something (Ungemah, 2024, p. 116). It isn't just individuals that collect. Communities and institutions start and maintain collections. Consider the book collections of libraries, the art and artifact collections of museums and galleries, the specimen collections of researchers and scientists, the data collection of tech companies and governments. Ungemah, quotes a professional organizer regarding the definition of a collection: "a collection is a group of specific things that are treasured, displayed, and have a natural ebb and flow, so that a collection is refined on an ongoing basis" (Ungemah, 2024, p. 117), however, this quote doesn't seem to capture collection-ness completely. The quote that the things have sentimental or emotional value and only hints at the aesthetic value of a collection in terms of its display and refinement; it doesn't say anything about a collection's ethical value, the context the collection might have been created in, or the means by which it was created.

When a collection (A) has elements that other collections of its kind (B) do not have there is a sense that collection *A* is a better collection than *B*. A collection of specimens is better if it is comprehensive. However, if it only contains common specimens, it isn't as good as a collection with a rarer element. The elements should all be distinctive in ways appropriate for their class, whether or not the element is rare in the world. For example, a comic book collector's collection is made better by having first edition prints of popular super-hero first appearances than one without them, whether or not it has many copies of easy to find comic books. The community of comic book collectors decide what items are worth pursuing or having, and this has to do not just with the formal properties of a comic book but the context of that comic book in relation to bigger story lines, the popularity of featured characters and so on. The set of properties determining the cultural value of a potential element is different for each class of collectible but makes sense for that class of collectible. For example, collectors of fine art paintings look for different attributes of paintings in the determination of quality of paintings-as-elements-of-collection than a collector of Nintendo Entertainment System (NES) Cartridges would look for in NES-cartridges-as-elements-of-collection.

A Valuable Collection

People, in general, value objects (physical or non-physical) beyond their monetary value because they imbue sentimental value onto those objects. This seems obvious to me, but in case you need convincing, I'll attempt a definition of sentimental value how it is determined. Before laying out a definition of 'Sentimental Value' I'll give a quick and dirty version of monetary value from Karl

Marx's theory of value and labor, and Aaron Ahuvia's explanation from "The Things We Love" to set up a contrast between the two types of value.

Marx says that something one makes for themselves is a product, and that product, if it has a social want associated with it, becomes a commodity. Commodities are affected by supply and demand, along with the labor-time of the producer, giving it a price or a monetary value. The difference between the product and commodity is the social want but what both have in common is that both require labor time (Marx, 1865). The social want is important in determining the commodity price. Under capitalism, however, needs and wants can be made artificial (or manufactured). A result of which is tricking the consumer into believing that their desires are innate. But for the individual producer of a non-commodity, social value is not important for determining sentimental value.

In *Fanatic,* Ungemah talks about a bottle collector, Dave, and his bottle collection.[5] For any popular[6] class of collection, competition for rare elements among collectors of that type is fierce, driving up the monetary value for collectors of that type of collectible element. As an example, Ungemah describes bottle collecting and rare bottles being sold for $200,000 at auctions. But finding these bottles involves digging through the earth or in middens–not as easy as just logging into eBay or Facebook Marketplace; an endeavor that helps attach sentimental value in the form of labor-spent. Here, the example shows sentimental value playing a role beyond rarity and difficulty of acquisition in its high price tag.

Ahuvia gives us a way to define the values imparting sentimental value qualitatively. In *The Things We Love* Ahuvia cites interviews he's conducted along with research done by others about how people relate to things, brands, and other people. In particular, for the owned things people value, Ahuvia writes about the relationships between those things and the people that own them (beyond the relation "X owns Y"). Ahuvia goes on to say that people value things because they relate to some other emotionally significant person,[7] calling these 'person-thing-person' relationships, where the thing stands in for the person. Ahuvia continues to say that sometimes the person's object relates to a time of their life or a particular experience they've had, calling these 'person-thing' relationships. Without getting too tied up in concepts of self-hood in a neuro-biological way, Ahuvia posits that the person who has the object expands their sense of self to include the other person and when the object becomes a stand in for the emotionally significant person, the object is included in that sense of self. For the *person-thing* relationship, the object is part of the sense of self as a stand-in for that time of the person's life. Additionally, and critically, Ahuvia says an object a person spends time working on will be valued more by that person than an object that has an anonymous provenance if those objects are of similar type and quality. Similar type means something like woolen-knit-sweaters-to woolen-knit-sweaters or lager-to-lager and not Empire-State-Buildings-to-scale-models-of-Empire-State-Buildings or Paintings-by-Jackson-Pollack-to-painting-by-a-5-year-old. Finally, liking is different from loving, which Ahuvia states changes the way a person would emotionally integrate these objects (Ahuvia, 2022).

Let's marry Marx and Ahuvia: Since there is a difference between liking and loving, it's reasonable to simply say the context of the relationship is a crucial factor in how value is imbued. Seeing your coworkers everyday hits differently from seeing your grandmother for the last time, which is different than from your feelings about time spent on a romantic relationship gone sour. Ahuvia shows us these qualitative experiences (and objects, eventually) distinguish themselves in our minds providing a way to construct a network of these relationships as they relate to us. The people that are friends with my best friend are, in some sense, vouched for so that there's a pre-knowing credit given, and their friends vouched for, and so on, down the line. This works for negatively sentimented relationships, too.[8] We also get from Ahuvia that objects can represent the person we have these relationships with, *including* ourselves. From Marx we get the idea that time spent laboring adds value in a quantitative way, but the idea of labor seems appropriate in the context of a product, and to include people and non-physical objects, I want to expand the definition of labor to be very general and something akin to 'time spent' or attention; Both imparting a quantitative value. For objects, it may simply be that the time spent on the object was spent yearning for it, like a teenager who has worked hundreds of hours to afford a used car, who then loves it, despite the car seeming to be junk to other people. The attentive time comes with imagining what it would be like to have the object and gaining knowledge of the object, or simply, familiarity as attention is paid and time is spent. The teenager dreams of what it would be like to have the car, and once possessed, cherishes it. In mathematical terminology we have the dimensions of sentiment from Ahuvia and the magnitude of sentiment from Marx. The vector it describes is sentimental value. As for edge cases, like sociopaths or the terminally greedy who only care about their own money, the (would be) personal relationships that act as edges to the nodes of people only are self-loops in the most extreme case; these people only relate to others via commerce.

So, sentimental value for a person relates something possessed by the person to another person that they care about, including themselves, in a sort of self-making. This self-making can be externalized by the possession or existence of physical objects, the presence of cared-for others, or may just be a dear memory. Sentimental value has real value to a person because that person considers it when making decisions regarding money. So, if money is said to have value then sentimental value also has value that can be converted into monetary value.

Before moving on from sentimental value I want to mention two more ways of assigning value to something; social value and cultural value. To paraphrase Marx social value is when *other* people value it. When we talk about organizations, communities, or institutions owning collections, this definition doesn't seem as satisfying. Pierre Bordieu, in *Distinction: A Social Critique of the Judgement of Taste* adds that those with cultural, social and economic capital are able to convert between them, as one of the advantages of privilege (Bourdieu, 1984). To be explicit, among collectors of some class those collectors who have in their collections rarer elements, or having a better collection will accrue social/cultural value in a way similar to the way that Bourdieu says the bourgeoisie would accumulate social/cultural capital.

In contrast, a painter creates a work of art that has sentimental value to them (or not, but I imagine one ideally finds their work intrinsically rewarding—and you definitely don't start as an artist for the pay). An artist's body of work attains cultural value as they work and become well known. That art exists in its particular cultural context in which it relates to other pieces of art, and in that context, shared meanings emerge. Beyond honing-of-craft, there may be a point where the artist's work accrues cultural/social value. Depending on this work's cultural/social value art collectors will pay in terms of economic value. If these were mere commodities, then the value added by labor is the same, but it is the cultural/social value that changes the economic value. An ideal art collector finds and distinguishes work by artists before the demanded economic value makes works by these artists unattainably expensive. These disparate kinds of value all play with or against each other, and each has its own distinguishing characteristics. For example, procuring of a rare or desirable element of a collection confers a cultural or social prestige among collectors of that class of element.

A Collection of Collections

People collect all kinds of things, from bottles to data. To distinguish from the broad use of "collecting" or "collection", I'll describe different sorts of things that could reasonably be called collections in English. From incidental collections to hoards to something between thes two extremes. I want to nail 'collection' down and talk about other things that are called collections which may or may not fit the collection definition given earlier.

The 'Incidental Collection' is a sort of collection but the category or theme of its contents marks a low bar. That bar might be something along the lines of "things in my home" or "things I use every day". Things in these "collections" might include flatware sets, socks, cotton swabs, and toothbrushes, and so on (these are our used-day-to-day possessions). There isn't any intent behind it beyond "I need a toothbrush" or "I like this art print". In one case, the object has instrumental value like: "I need a toothbrush so I can maintain oral hygiene." In another case, there is a sort of intrinsic value where you have it because it makes you happy or fits your preferences in some way: "I buy the thing—I use the thing—I own the thing." There isn't any great sentimental value associated in the toothbrush (or any other tool); they are simple things that fit into our lives when the need arises; things we manage not to rent or borrow from anyone we know. When we're done, they're dormant until the need arises again, or until we finally decide that we'd rather not think about it anymore or take care of it. At that point, we sell it, see if a friend needs one, or just throw it away.

In my home there is an incidental collection of books a visiting friend might see and think fits a penchant for collecting esoteric sci-fi novels and philosophy texts. My main motivation is to read something a library I had access to didn't have in its own collection or that I couldn't finish in three weeks. I don't have any special attachment to these things, generally, and the way we think about these sorts of things is similar to other instrumental household items. My spoons are valuable to me because they convey food from my bowl to my mouth. A book is valuable to me in the sense that I can read it, refer to it, look at the pictures, etc. My computer is valuable to me because of the capabilities it has that suit my preferences. But even though it is *my* computer it's a mass-produced object and immanently replaceable. Many of our things don't have any sentiment attached to them because most of our day-to-day possessions don't warrant it but we may store our possessions in a display-like fashion. Storing books upright on a bookshelf looks like display, but it makes them easier to access when you need a single book, as opposed to storing them in horizontally oriented stacks.[9] This is the case for most of our possessions. For an incidental collection, display seems only tangentially important compared to the instrumental accessibility or ease of use of its elements and is subordinate to its utility. Whether or not the items of the incidental collection are treasured seems to rest, at least for myself and other people I've talked to, on their sentimental value.

Other things in our home do have sentimental value: Pictures of our loved ones, souvenirs,[10] crafts we've made ourselves, gifts from others, and so on. I throw pottery in my spare time, and the bowls I use are special to me (i.e. they have sentimental value) because of my relationship with myself, via the bowls.[11] I could replace those bowls, and probably will, when I throw a set that suits my preferences and needs at that point in time. My relationship with myself via an object is different than with another person. Among the bowls I threw, one was sort of a freak of a bowl: a little bit lighter, and a little bit smaller, and not the same color of glazing. Yet, my partner refers to it as *her* bowl. When I mention that I intend to eventually replace those bowls, my partner resists the idea. The value of that bowl in particular and the bowls as a set aren't high in a monetary sense because they're definitely beginner level bowls, but the value to her is sentimental and a function of our relationship.[12] Either way, we both find it relatively easy to get rid of objects that we think are clutter. There isn't enough practical value or sentimental value to justify its continued use of space.

At the other end of the spectrum as a purely incidental collection, hoarders aren't making the same sentimental-instrumental valuations with their stuff that most of us are. Hoarding is an obsessive-compulsive mental disorder (Ungemah, 2024, p. 116) where someone has difficulty getting rid of things and it compromises their ability to live in their home.[13] Ungemah says that psychologists distinguish hoarding by three facets: the hoarder's lack of their use of the items they have, their inability to restrict those items to a space set aside for them, and the absence of time spent organizing and decluttering their possessions. The hoard is a hazard to the hoarder and the people in their lives (Ungemah, 2024, pp. 117,118). To further distinguish against what I will call a collection, the hoard in an aesthetic sense has a greater uniformity of elements (that is, repeated elements or elements that do not make the collection *better*), a lack of focus in terms of class or theme (unless the theme is the detritus of life and shopping) and an excess of potential instrumental value (everything can be used for some future project, or some future use).

Whereas "a collection is a group of specific things that are treasured, displayed, and have a natural ebb and flow, so [it]is refined on an ongoing basis," (Ungemah, 2024, p. 117), in a hoard, the value and/or quality of those items is low. Hoarders and collectors both have a compulsion toward the acquisition of new objects and both are emotionally attached to them, but a hoard is distinct from a collection because it is composed of low-material value stuff and, often, in bad condition.

Collections as non-Hoards and non-Incidental collections have themes or classes that span among almost anything you can think of (I'm hedging here; I don't know what you, dear reader, can think of) but collections need not be limited to physical objects. They can include experiences, like traveling, making art, playing video games,[14] or whatever. They can include electronic things, like data, or NFTs (still data, but that's unimportant). The list isn't meant to be exhaustive, but the *ideal collection* should be applicable to these things. Item can represent some sentiment. To push the idea of representation, consider a list of books one might keep to track what they've read, or plan to read. That list doesn't hold actual books but just references to individual books. The dearness of the list comes not from the list itself but, depending on its nature and number of copies, from a relationship between the reader, their past deeds, and future reads. A diary is a similar sort of externalization of experience. Both the reading list and the diary can relate the collector to themselves via the representations of things which represent the person themself. That is, something like a reading list represents the collection of experiences of reading each book to the reader, but is not itself the experience of reading any particular book. The item and collection simultaneously stand in for a person in relation to themselves. To push even further, imagine your long-passed father had a list of books that he had been reading over his life, and had planned to read a number of books but never got to them.[15] One level of abstraction says, "The book-list is important to me." The next level says, "The plans of my father are important to me." You find sentimental value in finishing the list for your father because that completion meant something *to him* and thus it does to you as well. Meanwhile, you find value in the list itself, regardless of the particulars of existence, because of this chain of caring, which ultimately ends with you, loving a person who exists in your head (as an abstract or mental object, take your pick). At one end of this chain is you and at the other end is also you. This explanation of sentiment towards non-physical objects can use the same mechanisms as sentiment for a physical object.

Groups of people, organizations, and institutions can also create collections. Think of library collections consisting not only of books but media of all sorts; with some (I'm familiar with the Octavia Lab at the LAPL Central Branch, and Burbank's library system also has one) (LAPL, 2026) including labs that house tools for members to use: from mundane tools like ladders and hammers to 3D-Printers, large format printers, digital illustration equipment, etc. (Kang, 2026). Seed banks are collections of seeds functioning as agricultural back up. Data Warehouses collect data and metadata[16] so that the owning organization can create reports, analyses, or predictions. Natural history museums have collections of animal specimens of all kinds, depending on the museum some may focus their collections on birds, sea life, or insects. Art museums house collections of art and depending on the size of the museum rotate their exhibitions of their collections and share collections with other museums.

Group or individual collections must also reckon with both access to their collections and the motivating value of those collections. If access is severely restricted, it becomes a stash, and if the motivating value is monetary it becomes something like investments or speculation. For collections owned by individuals composed of socially or culturally important elements, the individual has an ideal collection *if* the collection is based in sentiment or personal value. For example, a person who maintains a collection of religious artifacts of their faith has an ideal collection if they have a sentimental value for those elements of the collection. The group owned collection, though, needs only to have the cultural value associated with it.

SEEKING-COLLECTING

While there are sentimental and social reasons to start or maintain collections, the act of collecting itself is enjoyable (Ungemah, 2024, p. 108). For example, collections allow the collector to be part of a larger community who share their passion for the specific class of collectibles. Collectors of different objects form friendships and connections to others via their collections, through the internet, or at conventions, swap meets, etc. There are magazines that may talk about the collectible class so that object-finds (e.g. finding a rare bottle) or values can be discussed, or to instruct new collectors about the intricacies of that class of collection.

This social life surrounding collecting is rewarding in itself, since people value connection, and collecting as an act is as varied as the sought objects themselves. Ungemah's bottle collector, Dave, goes through a completely different process of finding bottles than someone buying those bottles, and Dave learns something about the community, or the people whose middens he excavates (Ungemah, 2024, p. 111). This process is different from a comic book collector searching through boxes at a comic bookstore, which includes its own form of digging (it's just in back issue boxes and not in a midden).

Darren Hudson Hick makes a claim about collecting, that, at least for comics, the search for missing issues to fill in the gaps (a "gappy experience") of a narrative has its own aesthetic value. The search for these comics engages something called the "SEEKING system" that in mammalian brains causes feelings of interest, curiosity, and anticipation. This system emphasizes that the act of searching for comics (or whatever collectible) is a reason to collect rather than solely having the collection itself. Hick declares "the pleasure of consuming the find temporarily replaces the pleasure of the hunt" (Hick, 2023, p. 684). Ungemah's Dave tells a similar story. For Dave, being on a dig is part of the thrill. To this end, Dave discards bottles that don't have any historical or personal value. Hick states that SEEKING helps to explain the high of, say, record collecting because that gappy experience can persist when a listener hasn't listened to a musician's entire oeuvre. The gappy experience does not set comic book collecting apart[17]—especially in 2026 when anyone can get a Marvel Unlimited subscription—but SEEKING can work to explain a lot of collecting-as-act, no matter the class of object.

As we've established, any object with sentimental value has that value because it is an artifact that represents a relationship to another person you care about. At a higher level of abstraction, since one way to reckon with our sense of ourselves or "who we think we are" is via what we do, the artifacts of our actions can be said to be artifacts of ourselves. James Delbourgo, in *A Noble Madness*, says while discussing Zhang Yanyuan, a 9th Century Chinese collector, "collecting is not something you do, it's who you are. A true collector is their collection" (Delbourgo, 2025, p. 44).

BUYING-COLLECTING

One it becomes a social activity, collecting, especially among strangers, involves exchanges of money for the item. Collections and collecting itself can also be influenced by monetary value, in the sense that someone with monetary capital can simply hire someone to collect on their behalf, or simply wait until the collectible pops up on a web-site or something similar. I term this 'buying-collecting'. Buying-collecting stems from exploiting a desire to collect, by claiming an underserved collectible status first, or using economic value to short circuit collection without the social, cultural, or sentimental value. Like Elon Musk pretending to be good at video games by paying others to play those games for him (MacDonald, 2025), a wealthy art collector could pay a dealer to find and buy art pieces for them.[18] It seems like, for the seeking-collector, the point of the collection is *doing* the act of collecting with the collection as a side effect; The monetary value of any particular element is irrelevant because the collection isn't for sale. However, the influence of economic value persists as a motivating force behind collecting, in that the wealthy buyer-collector prefers simply *having* the collection, the collecting activity is reduced to clicking "check out". Any individual object's possession stands in contrast to its utility value, interpreted very broadly to distinguish between art and artifacts, monuments, relics, and so on, at one end of a utility spectrum, and tools on the other. We see this in other ways: such as so-called collectible or special editions, or toys and knick-knacks marketed as collectible. This can take the form of real or manufactured rarity to create real or manufactured demand to collect, which we will call "market-collectability". Market collectability simply is a commodity marketed as collectible, with the producer designating its collectability, as opposed to those who would ostensibly collect it. In capitalism, recall that a consumer's needs and wants might not originate with them. This is one of the ways we end up with a simulacra of collecting.

Take, for example, a Kickstarter project. Kickstarter is a crowdfunding platform that takes advantage of this to remove the risk of producing a project. To incentivize people to 'back' a project 'creators' offer reward tiers that may include 'Kickstarter exclusives'. For board games it might be metal coins or premium components that are functionally identical,[19] but in some cases, it might be an expansion that is only available this way. The collectible status of these things is only conferred by the fact that before production has begun, the creator has decided to limit the quantities that are available. Manufactured scarcity is a marketing tactic used to manipulate potential buyers into a purchase that they wouldn't have otherwise made or to spend more money than they otherwise would have.[20]

There are more extreme cases where things are made for the sole purpose of being collected. Funko Pop! Vinyls are vinyl figurines that usually only have one point of articulation at the neck, but people collect them.[21] They are second-order collectibles (again, so-called), in that they use existing intellectual properties to create a collectible based on that. For example, Teenage Mutant Ninja Turtles (TMNT) was a comic book released in 1984 and later adapted for toys, video games and animation. Still a popular property, there are Funko Pops of the TMNT characters, but their collectability independent of TMNT seems flimsy, and any TMNT-ness is abstracted away from them by their Funko-pop-vinylization.

Funko Pop! vinyl figurines also have dubious aesthetic value with characters forced into a somewhat uniform shape and proportion. There doesn't seem to be anything artistically original in the Funko Pop! Vinyl case, since they reuse the same basic shape to represent pre-existing characters. The value of any collection of them is likewise dubious wtih figurines produced solely for their "collectability". There's no real scarcity since they're manufactured with collectability in mind, so any scarcity is also manufactured. In the case of a FOMO-mitigating add on for a Kickstarted board game, the content of the add on has some value in terms of adding game play functionality. In the Funko Pop! Vinyl case, there doesn't seem to be any value in acquiring a figurine from a limited run. The buyer, assuming it's a collectible, is likely to keep it in its box, because the motivation for collecting the item is that it retains monetary value.

There doesn't seem to be anything Funko does other than create figurines that use cute proportions and taking consumer money. Funko claims that "Everyone is a fan of something", but there doesn't seem to be a reason to buy a vinyl Spider-Man when you could just buy some more Spider-Man comic books, watch Spider-Man movies, or play Spider-Man video games. One might say, a true Spider-Fan would buy a vinyl Spider-Man, but buying a vinyl man does not make one a Spider-Man fan.

Theodore Adorno and Max Horkheimer, in "The Culture Industry: Enlightenment as Mass Deception", would agree with this. A shallow read of Adorno gives the impression like he doesn't like anything, but his claim (briefly) is that the consumer is passive in their role, and they consume the things they have access to, and those cultural artifacts and media are mass produced and their content and qualities narrow down to what most people will buy or consume, especially if they're re-hashes of intellectual properties that have already sold well (Horkheimer & Adorno, 2002). If a collection is made of mass produced objects, even if that production run is very limited, it is still made to maximize profit. Once people buy, the producer doesn't care if not everyone was able to get one because that would just cause more demand for the next production cycle. There is no reason the producer of an object (especially a mass produced one) should have any say that it's collectible or not since the communities doing the collecting should be the ones deciding or conferring the status of collectible.

A particularly strange version of collecting by way of manufactured scarcity is blind boxes or other secret-until-you-open-it things. When engaged in sealed play Magic (among other games and kinds of cards) booster packs work like this:[22] what you pull out of the pack is what you get to make a deck with. However, using an X-ray CT scanner, one is able to ‘peek’ at the contents of a sealed pack, and therefore know the contents while keeping packs in as new condition as possible (Wayne, 2024). The ability to subvert manufactured scarcity by ‘hacking’ into the pack shows a desire to find rare items while incurring a lesser per unit cost (of course, after the cost of access to a scanner) and passing the ‘duds’ to consumers who don’t know any better. The manufacturers aren’t losing anything by this, and the dealers are profiting from it, but in the end, the consumer who buys packs thinking they’re collecting wonders why they never pull anything great out of a pack.

Hick in a footnote says the focus on collectability almost destroyed the comic book industry. “Collectability” seems to be pervading and ruining other industries and collector’s realms. In the worst version of this Hasbro (who now owns Wizards of the Coast) seems to be “all in” on Universes Beyond, where other media properties are featured in Magic sets, such as The Walking Dead, or SpongeBob SquarePants. The initial reception was critical, because of a “dilution of Magic’s brand identity.” But sets that are well designed or thematically appropriate seem to be better received. The sets are viewed critically and the new players it draws don’t seem to have the cultural capital to be welcomed by existing Magic players.

Producers are exploiting the very human urge to be part of social groups, have hobbies, and just live their lives by offering products that can be made available without any essential limit. For Kickstarter creators, the risk of bringing something to market becomes negligible. In the case that the Kickstarter campaign wasn’t successfully funded, no one paid anything and everyone walks away. In the case that the Kickstarter campaign is funded, the creator finds enough demand and funding that there is no longer any risk of a product not selling, except through their own incompetence in planning or managing a project (or that of the manufacturers they contact, but causally, that’s still on the creator).[23] This might seem like a win for the consumer who may get a product suited to their needs and help bring it to market. Often, after the campaigns projects run behind schedule, or the creators were too ambitious, or world events affect shipping costs, or whatever. The consumer, consumed by FOMO, provides an interest free loan to the creator, who delivers on time *if* the consumer is lucky.[24] Hasbro likewise eliminates the risk of bringing a product to market by attempting to convert cultural or sentimental value to economic value, via their Universes Beyond releases. Producers of blind boxes and things like Labubu also attempt this conversion.[25]

For the act of collecting, it becomes unclear why people even begin to collect these objects. They are a simulacrum of collecting; they are completely divorced from the idea of putting any work in beyond waiting for the next drop, the next issue, the pre-release to release, or to get in on the lottery to buy the thing before the bots. Pick any of these from a myriad of ways to participate in the way to accrue cultural capital with your economic capital. For the objects themselves there's an analog for any of the re-prints or rereleases or so-called collector's items for the "true fan". None of these things are collectible until market interest as a collectible is whipped up (Baudrillard, 1994). If someone saw a Teenage Mutant Ninja Turtle Funk Pop! Vinyl, would they ever know that it was started as a comic book?

Jean Baudrilliard, in "Simulation and Simulacra", gives us a definition of simulacra that, via simulation, copies something, and through this process loses its link to the original. Eventually, this link is severed completely from reality, and the simulacra represents nothing (Baudrillard, 1994). Universes Beyond and Funko Pop! Vinyl are simulacra of collectibles, therefore the collectors of these things are simulacra collectors. There's always a new hotness, though, and what remains constant is the requirement that you pay for it. When a Simulacra-Collector can't get their collectible, they become indignant or angry (contrast this to Dave the bottle collector, who digs in the dirt of outhouse-middens without expectations and is thrilled with a find). The Simulacra-collector may make up stories about why they weren't able to get it–sometimes it is bots buying everything up, other times scalpers buy everything to resell it. The producer doesn't care about there are scalpers or that most of their customers can't get something; it's a system that works for them, because, again, producers work to manufacture extrinsic desires in consumers.

There isn't any ground for these collectors to be angry because the sought items were never anything they had a right to, and only become collectible because a manufacturer marketed them well enough to hoodwink people! The anger may stem from the sense they're losing cultural or social capital in their chosen group because that group says the new hotness is required, and by not being able to get it, they lose that value. Buying-collecting, because it roots itself in monetary value, as opposed to sentimental value, facilitates the generation of emotions regarded as negative (such as anger, disappointment, sadness, but I don't think they're negative per se) as opposed to those emotions regarded as positive (such as satisfaction, joy, accomplishment, but again, I think there's more to an emotion than what's on the box).[26] In the case of the Beanie Baby craze, there was a mania about "collecting" them and when the factors around the speculation (that is, the monetary value drive) died out, most people lost interest. Part of it was Ty Warner's deliberate manufactured scarcity, and part of it was his manufactured sentiment in the consumer, via the anthropomorphizing of them by giving them names, birthdates, etc. (Vintage Virtue, 2026). In 2026, there are still fans of Beanie Babies who have value for them beyond their monetary value. The mania around Beanie Babies began purely as a collectible from a market perspective pushing toward cultural and social value and further to sentimental value. For some, that sentiment exists, and for those whom it never did, they were collectors of nothing. Funko Pop! Vinyls capitalize on existing cultural mainstays and would seem to have more staying power than Beanie Babies because they understand that generating cultural and social value is harder than riding something else's value via licenses.

In addition to corrupting collection via manufactured scarcity, an outgrowth of buying-collecting is a move to subvert sentimental and cultural value to economic value by denying ownership of anything to people via licenses, leases, and memberships. One of the reasons people had for pushing back on Amazon also selling e-books and selling Kindles was that the e-book wasn't a physical object that you had possession of, and Amazon had the power to remove licenses to e-books from your Kindle.[27] Netflix used to send DVDs by mail until 2023, and now it's only streaming video. Cars are so unaffordable that dealers offer to lease them so that consumers can afford payments, without actual ownership. Jeff Bezos argues that people won't have computers in the future because they'll just use computing power on the internet—similar to the move from individually owned electrical generators to getting it as a public utility. We've seen a move toward that with streaming video game services like Amazon Luna, and Google Stadia (which Google shut down in 2023 (Google, 2026); luckily Google refunded any payments made for games or add-ons (9to5Google, 2022)). This is rent seeking pushed to the extreme. Private companies aren't utility companies, and while I argue that owning worthless knick-knacks is an imitation of collecting, having control of the things in your life also has value. Original movies can be 're-mastered' and changed, and digitally streamed, with no real announcement. This sort of thing made physical copies collectible, but now it can happen invisibly without recourse. Physical content in a collectors possession can't be changed, at least without the collector knowing about it.

Collectors of experiences go through this same process. The collectors responsible for their cabinet of curiosities and rapacious institutional museums 'collected' items from around the world. These objects were rare to European archaeologists, naturalists and travelers but taking these objects tore them from their context and reduced them to trinkets of other places. The owners of these collections vie for cultural status but the thematic class of these collections seems to have just been otherness. These aren't mere souvenirs—a token of an experience, but represent an attitude of entitlement. The attitude to "spend your money on experiences and not on things" is the modern version of this. With people achieving this through travel, but as a traveler the consumer is a tourist. With tourism a part of the economic machine, and so subject to the same manipulation of buying-collecting. The places popular to travel to are popular because of their cultural capital among the traveler's peers and community, but also among the places the society you live in says you should want to go to. The travel guides tell you what the best places to eat are or which are the best sites to see. The thing is everyone has their own tastes, and even if they don't, you don't need to fly across an ocean to eat at McDonald's. Everyone says "I'm going on an adventure!" but there aren't any dragons, only well-trod roads and higher prices if you can't speak the language. For a seeking-collecting experience, visit your family and friends, if you need a place to go. Visit it because you're curious. Visit the sites from your favorite tv shows or sites that inspired your favorite video games because even that is a sentimental attachment to a place. Visit a place because you love to travel, not because you think you're going to score "cool points". To wit, Atlas Obscura can tell you all the off-the-beaten path things for you to see and do, but even Atlas Obscura is trying to sell books and get clicks. If everything we do is seen through the lens of collection, or through *having* versus *doing*, everything is less valuable in a personal/sentimental sense, less valuable in a cultural value sense, and only produces value for the people taking your money.

CONCLUSION

Sentimental value is the foundational value for any genuine collection. Once it has sentimental value it is an ideal collection, and very often those collectors exist within a community of collectors of that collection kind. Cultural/social value also plays a part in community or group collections with those collections also exiting within a community of communities. Thirdly, money becomes a factor only after the procurement of an element requires it. It would be ridiculous for a library to purchase an expensive book for its collection, simply because it was expensive, and it is ridiculous for an individual collector to make a similar purchase.

The definition of sentimental value provides a way to explain the high prices that some collectible items command. Sentimental value was defined using Marx' theory of value, for products, and psychological research about people loving or valuing their possessions. Sentimental value, along with cultural and social value factor in determining market prices for potentially collectible elements. Cultural or social value is demonstrated when having certain elements of a collection can bring prestige to a collector and their collection.

For individuals, ideal collections or Seeking-collections are based in sentimental value and related to their collector-community with cultural or social value. There is a genuine quality to them. They are personally paid attention to with the collector usually displaying them, and are valuable in many facets, not just monetarily. They have an aesthetic quality to them brought out by their thematic class and the care of the elements.

Simulacra-collections or buying-collections seem superficially similar. The elements are valuable, but primarily via economic value, with little to no personal connection to the items. Their display is for a restricted group of people, if they are on display at all. The aesthetic qualities of the collection are diminished because of the economic motivator, which diffuses the focus of the thematic class. We see this with the cabinet of curiosities, and we see this in scalpers who buy up collectibles to resell.

Culture and social value are out of the control of any single person and are arbitrary because it only takes the agreement of enough people for something like a Beanie Baby to become collectable. This happens regardless of something being marketed as collectable by its producer. A numismatist (coin collector) could certainly collect uniform stacks of Sacajawea dollar coins (which are $1 per coin, in general) but a *better*, ideal, collection of coins would still pay attention to the elements composing the collection. Otherwise, it’s just so much simulacra.

[1] This becomes relevant later.
[2] Also known as MTG.
[3] These terms are interchangeable.
[4] Not restricted to the 21st century, but I am writing in the 21st century. To avoid losing focus for the paper, I'll just say it here; collecting, imperialism, and capitalism are all tied up with each other.
[5] Apparently, bottle collecting is the 3rd most common type of collecting, according to Ungemah.
[6] This is hand-wavy, but let's just say that it's popular if there are a lot of people collecting it, and some of them are strangers.
[7] This is not the only way that something can have value. Beyond sentimental or personal value, there is obviously economic or monetary value, but I will talk about social/cultural value. Beyond the scope of this paper, things can have religious value or some kind of mystic value, and instrumental value. Hopefully, it's obvious that this list isn't exhaustive, and some of these values may play a part in some collectability of something, like "Christian Relics" or something like that. For now, I have it divided along the lines of personal and extra-personal and simulated.
[8] This sounds like having a reputation, but I'm not thinking about it in the same way.
[9] I find that storing plates and like books is easier for me.
[10] This example will be discussed later, but for now it's important to recognize that there is sentimental value.
[11] Just in case you don't know, throwing pottery is to make pottery on a potter's wheel.
[12] I think this is why gift giving seems to be so difficult for some people. How do you appropriately gauge the correct non-monetary value of a gift such that it's not too intimate, and still intimate enough?
[13] Particularly in the West, but the discussion of hoarding and its interpretations is out of the scope of this paper.
[14] Live your life the way you want to.
[15] My father is still alive, and this is a bad example *for me* because he spends his retirement years binge watching science fiction movies and TV shows on streaming services, and he's still alive. Regardless I can imagine he was just tearing through those American Classics, and that I wrote this in 2050, and he had an unfinished list. I'm not going to go back to finish his streaming media quest though.
[16] Data about the data.
[17] You'd have to set rules for yourself in 2026 when you have access to pirated and legitimate copies of the work online, in a way that's like only playing with legitimate cards in a game of Magic, which you aren't strictly required to do.
[18] Henry Cavill is known for being an avid fan and collector of Warhammer and has the money to spend on it (Warhammer Community, 2022).
[19] I'll admit having splurged on the metal coins. But at least now I can talk like a pirate and about my love of the "clink" of coin.
[20] FOMO or Fear of Missing Out
[21] There may be Funko Pops with posable arms or some that aren't posable at all. I've only handled a dozen of them or so.
[22] You have to sell packs somehow, but I'm more okay with this since it's baked into the game's definition and has a use in game.
[23] It wouldn't be capitalism if Kickstarter didn't take a cut, which they do.
[24] Correct me if I'm wrong, but I thought capital justified itself because it was the one taking risks.
[25] Or Beanie Babies, or whatever's coming in a Post-Labubu world.
[26] I'm not as convinced that emotions are bad or good, compared to our reactions to our emotions being better or worse.
[27] To me, this seems related to the 'right to repair' that a lot of other technology companies dislike.

Bibliography

9to5Google. (2022, 11 9). *Stadia Refunds Start On November 9*. Retrieved from 9To5Google: https://9to5google.com/2022/11/09/stadia-refunds/

Ahuvia, A. (2022). *The Things We Love: How Our Passions Connect Us and Make Us Who We Are.* New York City: Little, Brown Spark.

Baudrillard, J. (1994). *Simulacra and Simulation.* (S. F. Glaser, Trans.) Ann Arbor: The University of Michigan Press.

Bourdieu, P. (1984). *Distinction: A Soicial Critique of the Judgement of Taste.* (R. Nice, Trans.) Cambridge, Massaschusetts: Harvard University Press.

Delbourgo, J. (2025). *A Noble Madness.* New York: W.W. Norton & Company, Inc.

Google. (2026, 3 20). *Stadia*. Retrieved from Stadia: https://stadia.google.com/gg/

Hick, D. H. (2023). Collecting for Closure. *Canadian Journal of Philosophy*, 678-687.

Horkheimer, M., & Adorno, T. (2002). The Culture Industry: Enlightenment as Mass Deception. In *Dialectic of Enlightenment* (B. o. University, Trans., pp. 94-136). Stanford, California: Stanford University Press.

Kang, H. (2026, 1 25). *Lending tool library set to open in Koreatown: 'Not everyone has space to store a 6-foot ladder'*. Retrieved from LAist: https://laist.com/news/lending-tool-library-set-to-open-in-koreatown

LAPL. (2026, 3 20). *LAPL Labs: Creativity Within Reach*. Retrieved from https://www.lapl.org/labs/octavia-lab: https://www.lapl.org/labs/octavia-lab

MacDonald, K. (2025, 1 25). Elon Musk Stands accused of pretending to be good at video games. *The Guardian*. Retrieved 3 20, 2026, from https://www.theguardian.com/games/2025/jan/20/elon-musk-stands-accused-of-pretending-to-be-good-at-video-games-the-irony-is-delicious

Marx, K. (1865, 6 27). *Value, Price and Profit.* Retrieved from Value, Price and Proft : https://www.marxists.org/archive/marx/works/download/pdf/value-price-profit.pdf

Ungemah, J. (2024). *Fanatic.* Oxford University Press.

Vintage Virtue. (2026, 3 20). *The Beanie Baby Craze: A Look Back at the 90s Collectible Phenomenon*. Retrieved from Vintage Virtue: https://vintagevirtue.net/blog/the-beanie-baby-craze-90s-collectible-phenomenon

Warhammer Community. (2022, 12 16). *Warhammer - Starring Henry Cavill*. Retrieved from Warhammer Community: https://www.warhammer-community.com/en-gb/articles/ZbMz2dj3/warhammer-starring-henry-cavill/

Wayne, A. (2024, July 16). *Youtube; Ahron Wayne's Strange Brain*. Retrieved from Youtube: https://www.youtube.com/watch?v=j7hkmrk63xc

FIRST ORDER CLAIMS IN MORALITY AND MATHEMATICS

Anthony Ulibarri

INTRODUCTION

One way to be a realist about morality is to maintain that there are nonnatural moral properties. Likewise, one way to be a realist about mathematics is to say that there are abstract mathematical (nonnatural) objects like numbers and sets. I call both positions versions of *nonnaturalist realism*. Mathematical *first-order* claims can be understood as conceptual truths and, as such, necessary, objective, truths. Cuneo and Shafer-Landau maintain that there are first order moral claims called "moral fixed points" that are indeed conceptual truths. *Moral fixed points* are intended to provide the moral nonnaturalist's position with more explanatory power and resources to respond to objections such as Evolutionary Debunking arguments, and supervenience challenges against nonnaturalism. In this paper, I (1) Defend moral fixed points, endorsing the thesis that there are first order moral claims that are conceptual truths and (2) reject the claim the conceptual truths like moral fixed points provide the moral nonnaturalist with more explanatory power, and argue that, instead, understanding mathematical and moral first order claims as conceptual truths undermines the core claims of nonnaturalist moral realism.

In Section 1, I give the background for moral and mathematical nonnaturalist realism. In Section 2, I argue, in agreement with Cuneo and Shafer-Landau, that many first order moral claims are best understood as conceptual truths, like objective mathematical truths. In Section 3, however, I argue further that understanding first order claims as conceptual truths does not give nonnaturalist realism more explanatory power because conceptual analysis is a human activity where nonnatural entities play no explanatory role. As an outcome of the preceding arguments, I conclude that prioritizing our first order commitments to these conceptual truths actually undermines the nonnaturalist realist's metaphysics.

SECTION 1: NONNATURAL REALISM

Realism in Morality and Mathematics, in the broadest sense, is the position that there are objective moral and mathematical truths, and that those truths are stance-independently or mind-independently true. The primary version of realism I'm concerned with in this paper is what I am calling 'nonnatural realism'.

The specific nonnatural moral realism I'll be assuming in this paper is the kind that holds that there are nonnatural entities like moral properties, such as *wrongness*. That means that this nonnatural moral realist maintains that these moral properties cannot be reduced to anything in the natural world, and so they are not the subject of empirical investigation by the natural sciences.[1] On the other hand, being a nonnaturalist realist about mathematics is a bit more nuanced but, I contend, just more of the same. I'll be focusing on the more commonly held view of mathematical Platonism. This kind of nonnaturalism posits nonnatural entities, such as numbers and sets, that are usually referred to as *abstract objects*. Since these abstract objects are non-spatiotemporal and non-causal, then they also aren't subject to the empirical investigation of the natural sciences, and hence they are nonnatural.

Now both views contain *first order* claims such as the following:

(A) "it is wrong to recreationally murder a fellow human being", and

(B) "no even number greater than 2 is prime".

So, for example, a nonnatural moral realist might claim that A is true because there is a nonnatural property *being wrong* that exists and which the act of 'recreational slaughtering of fellow humans' instantiates. In the same manner, the mathematical Platonist might claim there exists the number 2 and it is an abstract object which instantiates the property *being prime*, so B is true according to the mathematical Platonist[2]. More to the point, however, is that the first order claim B is *conceptually* true, based on our concept of natural numbers and the property of being prime, regardless of whether one believes numbers and properties are mind-independent objects or entities. So, consider the concept 'Prime', and the proposition that no even number greater than 2 is prime:

(**Prime**) A natural number $n > 1$ is prime iff n has only two positive divisors: 1 and itself.

(B*) <no even number greater than 2 is prime>

Based on our concept Prime, it follows that the natural number '3' (if it existed) satisfies, conceptually, the property of 'being prime'. The number 3 would just be prime, based on our concept of natural numbers and the concept Prime. If it so happens that some disturbed person approaches us and says, "well, to me, numbers go no higher than 2, so there is no 3 and it's not prime", then they aren't talking about the concept of natural number that mathematicians (and arguably everyone else) are talking about. It's fair to say that the discussion I want to have on nonnatural realism doesn't hinge on whether any of these first order claims are actually true (in the sense that there's actually an abstract object 3 out there, and it's prime). So, I'm not particularly concerned with whether our mathematical language actually intends to talk about abstract objects, or whether those intentions are true about reality. Rather, I'm concerned with whether they are conceptually true (in the sense that a sentence like B comes out of our concept of what a natural number is).

So, there are two important things we should take away from this background: (1) nonnatural realism says that there are nonnatural entities, like properties, numbers and sets, and (2) at least some of these claims are *conceptually* true. I've briefly covered here how being a nonnatural realist about mathematical claims meets these conditions.

Now the question is: Are there any first order *moral* claims that are that function in a similar way? The other question is whether facts about the objectivity of *conceptually* true first order claims provides nonnatural moral realism with more explanatory power. Ultimately, I reject the latter. But first, in the next section I explore and defend an argument by Cuneo and Shafer-Landau (2014) that there are first order moral claims called moral fixed points that are indeed conceptually true.

SECTION 2: DEFENDING MORAL FIXED POINTS

In Cuneo and Shafer-Landau (2014) the core claims that are relevant to the argument in this paper can be stated as the following:

> (C1) Moral fixed points are propositions that concern those situations that agents find themselves in and that are subject to or invoke moral judgement and assessment.

(C2) A system M counts as a moral system only if a range of moral fixed points are included in M.

(C3) Moral fixed points are conceptual truths.

In the next three sections I will explain and defend C1 – C3.

2.1 What Moral Fixed Points Are

The broad idea is that moral fixed points as defined in C1 are propositions for sentences like A, "it is wrong to recreationally murder a fellow human being"[3]. The main idea is that these propositions contain concepts that essentially have moral valence. Think of the situation of concern in A involving the *concepts* of 'recreational slaughter' or 'murder'. Pre-theoretically, (to me at least), any attempt at conceptualizing 'recreational slaughter of other humans' for any actions that happen to fall under the concept (for the person conceptualizing) are subject to negative moral judgment and assessment. Theoretically speaking, our philosophical theorizing might say that in all near possible worlds, where humans are like us, the situation found in A is always bad, to say the least. We might want to stretch this theorizing a bit more, however, to get at why that is.

Although Cuneo and Shafer-Landau do not make this distinction specifically, the idea here is that a proposition like A has nothing to do with sociocultural appraisals of harm, death, violence, or other such token acts that may or may not fall under the category of senseless murder or slaughter for said social culture. Propositions like A are concerned with the *concepts* used to describe acts that are, by definition, negative in their conceptualization. That is, wrongness is what it is to be the *concepts* of 'murder' in a similar way that 'being unmarried' is what it is to be a 'bachelor'. Bengson, Cuneo, and Shafer-Landau express the idea in the following:

> "…the concept [*recreationally slaughtering fellow persons*] is non-normative: it is neither an evaluative, directive, fittingness, favoring, nor aretaic concept. At the same time, the [*concept*] [*recreationally slaughtering fellow persons*] is not accidentally related to the moral concept [*wrong*]. What it is to be the concept [*recreationally slaughtering fellow persons*] is to be a concept that applies to an action if that action also falls under the deontic concept with a negative moral valence", (Bengson et al, 2024, p.99).

So, either a group or community has these concepts, or they don't. For a community to lack these concepts is not the same as a community maintaining the first order claim that 'recreational murder of fellow humans is good'. The former might be possible but the latter, I claim, is not possible. For example, the claim that Roman or Mesoamerican cultures participated in or observed ritual acts of violence is one thing, but those historical realities are not directly related to whatever *concept* they deemed to be "recreational murder", in addition to whichever members of the human race fall under their concept of human/fellow persons in a given context. Whatever the situation, these cultures, like all others, have these *concepts,* for whatever words or situations they use or have them in. They either had cultural reasons for performing such acts we now consider recreational murder, and they would not have described them conceptually as 'murder', or they committed these acts anyways with the knowledge that they were bad, i.e. justifying killing civilians in warfare for no reason, yet there's no reason to reject they still had the concepts for recreational murder. Even if they considered such acts justified, it would be quite a stretch to imagine that the *concepts* for recreational murder had positive or neutral moral valence, suggesting that they would positively or neutrally valuate these acts if they were done to themselves or loved ones. So, for any culture, whatever category acts of violence, (i.e. coliseum games, religious sacrifice, slaughter of civilians in war, ritualistic violence), fall into, they either do or do not fall under the *concept* of murder, on their own terms, yet the concepts still existed. It's perfectly consistent that they performed or participated in these acts and still had the concepts of "recreational" and "murder" or to describe acts they either didn't want themselves or loved ones to suffer through. The purpose of moral fixed points is just to point out that *everyone* has these *concepts*, however thin or minimal we consider them relative to the present day. So, if there's been any moral change or progress relative to time and place, it's due to the fact that more 'acts' have simply fallen under the concepts of 'murder' or 'recreational slaughter'. The *concepts* themselves have not changed, it's just that more acts and situations are falling under the concepts.

Now one might argue that it's possible that murder isn't wrong in some possible world. Perhaps, some possible world exists where murder is a virtue or enjoyed with bliss and excitement. My response is that this explanation is not the concept 'murder'. Our number theory skeptic above falls into the same category. Their claim is that B is conceptually false because there is no successor of 2. But their concepts of natural number and a successor function S, even if considered consistent in a model, are not associated with what *we* mean by number. So, in the same way, a basic number theory or arithmetic also has first order claims that are fixed points. The skeptic's challenge is not against the set of natural numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots n\}$ and structure $(\mathbb{N}, 0, S)$ for their argument utilizes the set $\mathbb{N}^* = \{0^*, 1^*, 2^*\}$, in conjunction with their own successor function of S^*, $(\mathbb{N}^*, 0^*, S)$. Likewise, the mathematical fixed points don't change, and so the first order claims that are candidates for these fixed points don't change either. For, any world where "3* is not prime" is true, 3* is not a number for our concept of natural number. Similarly, for someone in another possible world perhaps murder* (their concept) is not bad, but murder (our concept) is bad at both the actual world and any other possible world because that's just what the concept means[4].

2.2 Framing a Moral System

C2 is the claim that whenever we think of a system M as a moral system, we are thinking about the moral fixed points, regardless of how few humans these moral fixed points apply to. I think this premise sounds more controversial than it actually is. Recall, from the previous section, the moral fixed points have little do with whether these concepts apply to all humans, or which token situations fall under the concepts in a given context. It also has little to do with whether these acts are ever justified for any given reason. What is involved here is that anything we consider a minimal moral system is going to have as its constituent, a moral fixed point like A, that recreational murder is *wrong* because that's just what the concepts in A fall under for any concept of wrong. Cuneo and Shafer-Landau (2014) express that moral fixed points constitute the "boundaries" of any "minimally eccentric moral system" M that is to be distinguished from other normative systems.

In order to frame C2 a little better, it's helpful to start thinking about how the priority of some moral first order claims motivate our metaethical commitments. In an argument against nonnatural moral realism Hayward (2019) challenges the moral status of believing in nonnatural facts. The argument goes that it is morally objectionable to conditionalize our first order commitments on whether there are nonnatural facts (or nonnatural properties for the sake of my argument). The objectionable part is that it might turn out that there is no nonnatural moral property that recreational murder instantiates, thus it turns out that it's not wrong. The way I see it, the right response here is like that endorsed in Enoch (2020), a proponent of a robust nonnatural moral realism, that if this were the case, then it would be shocking enough to abandon one's commitment to this version of realism. What's important, in my view, is that the first order claims I'm talking about here take this kind of priority. So much so that, I argue, abandoning nonnatural moral realism for the above reasons is the same as not recognizing nonnatural moral realism as a candidate for a valid moral system[5]. For any attempt at conceptualizing a moral system that does not account for and explain our experience and phenomenology associated with the concept recreational murder is not a valid moral system, since those experiences just have negative moral valence. Another way to put this is to imagine using the same concept but with positive or neutral moral valence, such as 'recreational murder of humans is good' or 'recreational murder of humans is neither good nor bad'. If the concept of 'recreational murder of humans' is positive or neutral in a moral system (hence it is positively normative or non-normative) then such a system doesn't reflect the constituent concepts of 'recreational murder and 'human', thus it is not a valid moral system for us as humans. Perhaps there are some first order claims that don't merit this kind of reaction, but definitely not the moral fixed points that take priority, as defended here.

Bengson, Cuneo and Shafer-Landau (2024) also make a case that a metaethical theory should, at minimum, accommodate and explain the data that we consider as describing the moral domain. It's quite possible that the datum for a first order claim involving specific sexual conventions and norms, for example 'Premarital sex is wrong', do not accurately describe the moral domain. So, then it's possible that if a system does not accommodate this datum that it is still a moral system. That premarital sex is wrong might be true, but nothing in a conjunctive concept 'premarital' and 'sex' is such that it essentially has negative moral valence. If my argument so far is cogent, however, it's *not* possible that the datum of a first order moral claim involving something like recreational murder does not accurately describe the moral domain, because the meanings of those concepts are inextricably moral in nature[6]. Some first order moral claims take priority, such as the moral fixed points being advocated for here. And this is applicable to other metaethical theories, not just realism. Forms of constructivism, metaethical relativism, error theory, and expressivism can all be structured in a way that accounts for these primary first order moral claims so as to make them valid candidates for moral systems[7].

I think similar claims hold for mathematics and set theory. Take our skeptic from the previous section. It can be argued (I think definitively) that whatever structure $(\mathbb{N}^*, 0^*, S)$ is, it is not a foundational mathematical structure. The set concept $\mathbb{N}^* = \{0^*, 1^*\}$ might be a real concept, but just not a mathematical one. For the concept $\mathbb{N}^* = \{0^*, 1^*\}$ does not represent the set of natural numbers in any system that we consider to be mathematical, so any structure that contains $\mathbb{N}^* = \{0^*, 1^*\}$ as foundational is not a mathematical system, as opposed to $\mathbb{N}$. The important point here is that not all first order claims have this kind of status. If you consider the Continuum Hypothesis (CH), which states that there is no set between the infinite sets of the integers and reals, it's possible that two independent set theories, where CH is true in one and false in the other, are both consistent. That is, we can have one notion of set where CH is true, and another notion of set where ~CH is true. An idea that supports the view that both CH and ~CH concepts of set exist is found in pluralist conceptions of mathematics. Versions of this view can be found in, but certainly not limited to, Balaguer (1998, 2009), Hamkins (2012), Jonas (2020, 2024), and Zalta (2024). The idea here is that there is no single universe of set in mathematical reality, but rather a plurality of universes in mathematical reality, where some first order claims are true in a universe and false in another. What that means is that CH is independent of set theoretic axioms, such that you can take it or leave it when it comes to your set theoretic intuitions. Another way to put it is that two people can independently have intuitions about legitimate conceptions of collections or sets, one in which CH is true, and another in which it's false, Potter (2004), Hamkins (2012). Mathematical pluralism aside, if one were to somehow find indisputable evidence that there is no set between the integers and the reals that exists in platonic reality, so that CH is false, then this would be far from shocking enough to give up on one's realist view of sets (that sets actually exist in platonic reality, and there's no set between the integers and the reals in platonic reality)[8]. You would just be confident enough to say that CH is false, considering your updated notion of set + ~CH foundational to mathematics, and go on with your day.

2.3 Moral Fixed Points as Conceptual Truths

If the reader finds that they agree with the previous two sections, or finds them non-controversial, then it shouldn't be surprising that I argue here that the prioritized first order moral claims referenced above are just conceptual truths. So, conceptual truths like "recreational murder is wrong", either is or is derived from prioritized first order claims (also conceptual truths) that mark the boundaries of whatever system we're distinguishing. The reason why these first order claims work for such demarcation is because they are precisely what we mean by the words or concepts being used. The moral fixed points are just precisely what we mean using the terms or concepts of *wrong* or *being* (*im*)*moral*, regardless of whether these principles apply to all humans at all times.

I think it's best to argue this by making a clarification and by making an illustration. When I say that the moral fixed points are what we mean by *wrongness* or the property *being moral*, I'm not talking about which moral system or concept of morality or wrongness the words "morality" and "wrong" actually refer to. I'm not making the claim that when people use the word "wrong" they are using it to refer to, for example, the realist's conception of wrongness, or the expressivist's conception of wrongness. It's possible that both, one of, or neither of these systems are being referred to by people when they use the words "wrong" or "moral". Ultimately, I believe, that is an empirical scientific question for, at least, psychology and the neurosciences. What I mean is that I'm concerned with the kinds of things, subjects, situations, facts, etc., the words wrongness and morality are about.

Say that an extraterrestrial visits Earth and is trying to understand our customs and language. The alien observes our customs and is trying to map what they can to our language and asks what we *mean* by "wrong" or "moral". Say that they're talking to a philosopher, and the philosopher answers that what we *mean* is either the concepts of moral realism, or moral relativism or moral expressivism, etc., then they proceed to explain those moral concepts and the relevant systems. The alien thinks for a second and says that that answer fails to address their question (but they can find out if we want by attaching our brains to some apparatus that will map our psychological and neural activity to the corresponding moral concepts when we use those words). What the alien really wants to know is what kinds of things, or situations fall under the concept of being wrong or being moral/immoral, and what are the concepts of those situations. I think the simple answer for the alien is to draw their attention to a senseless crime, a humanitarian crisis, a mass shooting, and say "that, that's what I mean by wrong". We're going to point them to what, I argue, are those moral fixed points because that's just what we *mean* by the terms wrongness and the moral domain.

The philosophical reader might be thinking, "well no, the concept *wrong* is the philosophical moral concept x, or y, or z based one our usage and discourse". But that's conflating the question with 'what is the correct concept for our usage and conventions?'. For instance, it's not part of the concept 'number' that a number is an abstract object. We might find that out our usage of mathematical language and how mathematical sentences are made true corresponds to a metaphysical concept of what a number *is,* consequentially *that* metaphysical concept is the correct one. However, the concept 'number' is just the mathematical foundations, axioms and structure that number discourse and discrete magnitude is about (or possibly even less than that). No one is learning to solve equations by learning that numbers are abstract objects. In the same way, the concept 'morality' or 'wrong' is not 'mind-independent nonnatural property'. In the same way, if someone who has never seen water before asks us about the concept of water, we shouldn't say that it's H_2O. Rather we want to say the concept is 'clear liquid stuff that is safe to drink' or something to that extent. In the same way, that wrongness is a mind-independent nonnatural property might be true, but no one is learning what moral situations are by learning that morals are nonnatural properties. Because the conceptual truths are just what we mean by wrong, not the metaphysical theses that might support the concepts.

SECTION 3: CONCEPTS, NONNATURAL REALISM, AND MORAL FIXED POINTS

In the previous section I argued that there are first order moral claims that are conceptual truths, similar to the kinds of conceptual truths we find in mathematics. Cuneo and Shafer-Landau (2014) call these moral claims moral fixed points. The moral fixed points are also supposed to defend and give more explanatory power to nonnatural moral realism. However, although I defend moral fixed points, below I argue that in fact the moral fixed points play no explanatory role in nonnatural realism's main theses and that they neither strengthen nor defend the position.

3.1 A VIEW ON CONCEPTS

Up to this point I have been discussing concepts and conceptual truths. Now, I want to cover what concepts are and where they come from. Unsurprisingly, you can be a nonnaturalist realist about concepts, too. That is, a view where concepts are just abstract objects, nonnatural entities like mathematical objects or nonnatural moral properties. Also, up to this point, I haven't taken a stand on whether nonnatural entities exist or not. On my view, it doesn't really matter. I think this is important because what I find are the most compelling reasons that motivate the argument that moral fixed points give nonnatural moral realism more explanatory power, are the facts about how our concepts somehow depend on nonnnatural entities. If facts about abstract objects don't have relevance to our concepts, then there's no reason to think that conceptual truths like moral fixed points give nonnatural moral realism more explanatory power. So, my argument goes.

In Balaguer (2021) we find a view on what settles our conceptual analysis questions. It can be explained as follows. If concepts exist in the platonic realm as abstract objects, then they all exist. For example, there's concepts for our notion of the natural number structure, as well as some other concept for a natural number* structure (the structure ($\mathbb{N}^*$, 0^*, S)) like our mathematical skeptic above. But which one is the right one, and what facts determine and settle which one is the right one? In other words, when we talk about our concept of mathematics, what determines whether our foundations refer to the concept natural numbers or natural numbers* and their relevant properties? The answer, according to this view, and the view I'm taking, are facts about us, and our own usage and intentions (ibid, 2021). In the same way, what determines whether murder or murder* fall under the concept of wrongness are facts about us, and our usage and intentions. So, the combined concept of *recreational murder* just falls under the concept *wrong* because that's what we mean by *wrong* based on our usage and intentions.

A common response is that only the correct nonnatural entity (for example, natural number, not natural number*) exists, so our concept is just the one that exists in the platonic realm. However, recall that suggesting nonnatural entities in the platonic realm is an ontological thesis. One is committed to the existence of mind/stance-independent, non-spatiotemporal abstract objects or nonnatural properties. And arguing that one nonnatural entity exists as opposed to another is an arbitrary ontological commitment. One would have to argue for non-arbitrary reasons that one abstract object exists, and another doesn't, much like one would have to argue that the nonnatural property red exists but not blue, thus we should give more precedence to red as a color concept because red instantiates a nonnatural property and blue does not. But they're both color concepts and we have no better reason for saying the property for red exists over blue. The reasonable conclusion is that all concepts exist and the facts that determine which properties fall under the concept color are our own usage and intentions.

3.2 NONNATURAL REALISM AND CONCEPTS

I argue that the position on concepts found in the previous section is correct. And this would be no direct rejection of the ontological thesis of nonnatural realism. In fact, the simple version of the argument is that it really doesn't matter at all whether these nonnatural entities exist, because our concepts come from us and our mental activity. Of course, it might be the case that our usage and intentions, and mental activity, is being determined from other factors in reality, such as natural properties (i.e. our evolved biological structures, physical properties of our environments, etc.). However, any suggestion of this thesis just abandons the need or explanatory power of nonnatural realism, since if conceptual analysis questions are determined by us + natural properties, then there is no requirement for a nonnaturalist position.

That said, it can be argued (defensibly) that our intuitions sometimes lead us to think that mathematical first-order claims are somehow stronger or are grounded or more metaphysically advantaged than moral first-order claims. Clarke-Doane (2020) argues that mathematical first-order claims are no more metaphysically advantaged than moral first-order claims for the realist. I agree, and to go further, if the above analysis on concepts is correct, then mathematical concepts are also determined and settled, not by nonnatural entities, but us, and our minds, (and possibly natural properties). Questions about mathematical structures like natural number structures can be answered based on the structures we are intending to describe. So, what we mean by natural number is just the structure we intend to describe, Balaguer (2009). The fact that there are conceptual constraints for what counts as a foundational structure for our concept mathematics does nothing to explain the existence of mathematical abstract objects.

3.3 Moral Fixed Points and Conceptual Truth

What all this means for moral fixed points is that the fact they can be described as conceptual truths does absolutely no favors for the explanatory power of moral nonnatural realism. Like mathematical concepts, the fact that there are conceptual constraints on what counts as morality does nothing to further explain that there is a single correct concept for wrongness that nonnatural properties fall under. What that fact explains is how conceptual analysis is determined by our usage and intentions and our own mental activity.

To gauge the results, consider the kinds of evolutionary debunking arguments like those endorsed by Street (2006), and Cuneo and Shafer-Landau's response to such arguments from moral fixed points, Cuneo and Shafer-Landau (2014). Summarizing a debunking argument against nonnaturalism can go like this:

> 1. There are indefinitely many moral systems that can be conceptualized, all of which are possible.
>
> 2. Evolutionary forces have caused us to endorse only a small subset of all possible systems.
>
> 3. If moral nonnaturalism is true, then there is a uniquely correct moral system of stance-independent moral truths.
>
> 4. Given the vast range of conceptually possible moral systems, the odds that evolutionary forces have pushed us to endorse the uniquely correct moral system of stance-independent moral truths are extremely low.
>
> 5. Such odds entail that if moral nonnaturalism is true, then it would be a remarkable coincidence were our moral beliefs largely on target

The response from moral fixed points (from section 2.2), and which I have argued for at length, is that premise 1 is false. Since moral fixed points set the boundaries on what counts as a moral system, then it turns out that there are conceptual constraints for what counts as a moral system. Cuneo and Shafer-Landau, however, attempt to represent nonnaturalism in such a way that takes a step further, arguing that these conceptual constraints "guarantee" that we pick out a version of the uniquely correct moral system in premise 3 (ibid, 2014). But if we advocate the above position on concepts, then no facts about nonnatural moral concepts or nonnatural moral properties determine this conceptual constraint. Our usage and intentions determine the conceptual constraints. So, ironically, since I've argued at length for moral fixed points, I agree that premise 1 is false, and that the argument doesn't work. Yet, even if I reject premise 5, the only reason there is no coincidence is not because of any facts about nonnatural entities. It's only because of facts about us and our own minds. We determine what falls under the concept of wrong, and things like the moral fixed points are just those things we consider when conceptualizing morality.

Another argument is what I generically call the Supervenience argument. Versions of this argument have appeared in different forms, notably Balaguer (1998) and McPherson (2012). Since the argument applies to an array of metaphysical questions, I will summarize and narrow it down for relevance.

> (SUP): Properties A and B are discontinuous, if A and B are neither reducible to one another, identical to one another, nor metaphysically continuous with each other. Nonnatural entitites are non-spatiotemporal. If A is a nonnatural property and B is a natural property, and if the metaphysics of nonnaturalism is true, then if A supervenes on B, then the connection between A and B is brute.

To apply SUP, say there are two possible worlds w and w'. In w the metaphysics of a robust nonnaturalism are true (there are abstract objects, nonnatural moral properties and entities). In w' the metaphysics of nonnaturalism are false (there are no abstract objects, no nonnatural moral properties or entities). Now, imagine you suddenly appear in w', having spent your whole life in w. In w, you delight in the fact that mathematical claims like "3 is prime" are true, since all the abstract objects that make this true exist. Now you're in w', this same claim is now false, and you suddenly begin experiencing a life changing existential crisis where all your mental faculties can no longer conceptualize any mathematical facts at all. This all is all meant to be satire, of course. But, more seriously, one has limited options when explaining why this would or would not happen. If it were the case that you suffer no existential crisis and your mental faculties can still conceptualize mathematical claims, then it's hard to imagine what the difference would be between w and w'. Of course, the claim "3 is prime" would technically be false in w', however your experience of the magnitude physical properties in the world suffers no impact and your equations carry on as normal. Yet, if you deny this, and if there were major impacts on your mental faculties and experience of the world in w', because A supervenes on B then, per SUP, the connection that causes this would have to be a brute connection, since you would have to explain this kind of relation between two discontinuous properties.

Again, moral fixed points are meant to explain this connection. The connection is not brute because, according to the argument from moral fixed points, the connection that obtains between moral and natural properties is by way of conceptual necessity. And this connection, accordingly, would be grounded in the metaphysics of the nonnatural, since the connection is grounded in the metaphysics of concepts. But the same previous reply is effective here. Even if all of this is true, whatever connection is involved in the metaphysics of concepts, the connection is simply to our own usage and intentions. It's also unclear how this response from moral fixed points benefits moral nonnaturalism as it pertains to what non-spatiotemporal concepts are doing for our conceptual faculties. If SUP is true, and we adopt the view on concepts from 3.1, then it also applies to the abstract objects of concepts, and bruteness is still a problem for conceptual analysis. In the end, although I think moral fixed points make great strides to explain what morality can possibly be, they do nothing to commit one to a unique concept of morality itself, and they give no greater explanatory power to moral nonnaturalism as a concept of morality.

3.4 The Priority of Our Fixed Points

To finish section 3, I want to draw attention to a point I find interesting. What I find intriguing is that the kind of objectivity necessary truths like fixed points offer, I've argued, do nothing for nonnaturalist realism to gain more explanatory power. In fact, the greater the priority we grant these fixed points, the significance of this kind of objectivity grants no benefit to the nonnaturalist realist. In Zormpalas (2025), we find an endorsement for a principle of revisability to our prioritized first-order commitments. Zormpalas argues that sometimes our first-order claims are indeed conditionalized to our metaethical views. For example, say you abstain from eating oysters because you believe that oysters are sentient and capable of suffering, and you have some metaethical view where the properties of sentience and suffering and some other actions, instantiate some moral property. Later on, though, a marine researcher friend of yours tells you that they are now confident that oysters are not sentient and are not capable of suffering. Do you change your first-order commitment of not eating oysters? I don't think this will happen. For if you do change this first-order commitment, then you really just have another concept that falls under your concept of wrong. You might indefinitely be committed to abstaining from eating *oysters* (the marine bivalve mollusks that are sentient and capable of suffering), but you might have no problem eating *oysters** (the marine bivalve mollusks that are not sentient and not capable of suffering). It just turns out that only oysters* exist and not oysters. Just like number and number*, or murder and murder*, we have different concepts being considered. Now, I'm not sure whether the concept *sentient-creature-capable-of-suffering* in conjunction with a concept of harm is robust enough to necessarily fall under a concept of wrongness, but I've argued that the concept of *recreational-murder-of-humans* does. It can be argued that the greater priority should be given to the moral fixed point concept of recreational murder of humans (unless one thinks eating oysters is also murder, which I will not take a side on here). But the possibly less prioritized problem of eating oysters is still objective according to my explanation, and the prioritized moral fixed point is arguably objective in a way that first-order claims about eating oysters is not. But if previous arguments are sound, then no increase in objectivity does anything for the explanatory power of nonnatural realism. The objectivity might come from the necessary truths of the fixed points, yet, the necessary truths come from the concepts, which just come from us.

Conclusion

In this paper, I've argued that there are both mathematical and moral first-order claims that can be understood as conceptual truths, while considering nonnatural realism. I have advocated for a particular view of some moral first-order claims, called moral fixed points. The moral fixed points were supposed to offer more explanatory power against arguments directed towards nonnatural moral realism; however, I concluded that the opposite is the case. After evaluating some mathematical first-order claims and the moral fixed points, I've found that the analogies are similar between them, and that the concepts we find in these claims ultimately come from our own usage and intentions, and that the nonnatural entities found in the metaphysics of nonnatural realism play no substantial role in our conceptualizations. In fact, if nonnatural realism offers any kind of claim to objectivity I find that the objectivity just comes from the necessary truths of the subject first-order claims, and that what is necessarily true are the concepts that, again, come from our own usage, intentions and mental activity.

END NOTES

[1] Moral Naturalism on the other hand maintains that moral properties are natural properties, making these properties, in principle, subject to empirical investigation. It is possible that one can believe that there are both nonnatural moral properties and natural moral properties. It is also possible for one to hold that there are simply nonnatural moral *truths* and not nonnatural moral properties.
[2] The mathematical realist might also just claim that only structures or sets representative of the number 2 exist.
[3] In Cuneo and Shafer-Landau (2014) and (2024) they use the word "slaughter" instead of "murder". I take both terms to mean something close to the concept of "killing without consent", however, to avoid any ambiguities I use the less ambiguous term "murder". Although, I personally think the word "slaughter" already implies the sort of non-consensual violence and cruelty not implied by the generic term "kill".
[4] I suppose one can argue that the properties instantiated by P are not essential to the concept of 'murder', in the sense that the property 'being such that my head is severed' is not essential to the concept 'murder'. In this case, only the phenomenology and experiences P are essential to the concept 'murder', etc., and only P essentially has negative moral valence. That seems possible to me, (that the physical and phenomenological properties instantiated in these concepts come apart), but my focus here also includes the concept 'human', so my concern is what humans think murder is.
[5] This is interesting, because a hypothetical case like this can sustain the belief in a moral realist ontology, yet also claim that such a reality just fails to capture our human conception of morality, just part of it. Such a position can maintain that nonnaturalist realism, for example, isn't enough to explain morality but rather nonnaturalism + M*, where M* covers the rest of our moral concepts via some other means. While this situation is strange and purely hypothetical, it doesn't seem contradictory to me. A fictionalist approach to morality might make sense for some moral claims in this case, while being a moral realist about other moral claims.
[6] I am not asserting here that sexual norms or conventions are not considered the subject of the moral domain for some people. What I am saying here is that it is quite conceivable that there are no moral properties for sexual norms and conventions, and all the other properties describe what we mean by morality. However, I argue, it is *not* conceivable that there are no moral properties for something like recreational murder of humans and that all other moral properties comprise everything we mean by 'morality'.
[7] I have no intention to defend this point here, but it's clear that Bengson, Cuneo and Shafer-Landau (2024), all nonnatural moral realists, support the idea that metaethical theories opposed to nonnaturalism can be framed and constructed in such a way as to accommodate and explain at least some of the important metaethical data (they would challenge these moral systems on other grounds)—which is a good enough defense for the purposes of this section.
[8] Additionally, one could be a platonic realist about set theory + ~CH, and a fictionalist about set theory + CH, and mathematical practice would proceed business usual.

Bibliography

1. Balaguer, Mark. 1998. Platonism and Anti-Platonism in Mathematics. Oxford University Press, Incorporated.

2. Balaguer, M. 2009. "Fictionalism, Theft, and the Story of Mathematics." Philosophia Mathematica 17 (2): 131–62.

3. Balaguer, Mark. 2021. Metaphysics, Sophistry, and Illusion: Toward a Widespread Non-Factualism. Oxford University Press.

4. Bengson, John, Terence Cuneo, and Russ Shafer-Landau. 2024. The Moral Universe. Oxford university press.

5. Clarke-Doane, Justin. 2020. Morality and Mathematics. Oxford University Press USA - OSO.

6. Cuneo, Terence, and Russ Shafer-Landau. 2014. "The Moral Fixed Points: New Directions for Moral Nonnaturalism." Philosophical Studies 171 (3): 399–443.

7. Enoch, David. 2021. "Thanks, We're Good: Why Moral Realism Is Not Morally Objectionable." Philosophical Studies 178 (5): 1689–99.

8. Hamkins, Joel David. 2012. "The Set-Theoretic Multiverse." The Review of Symbolic Logic 5 (3): 416–49.

9. Hayward, Max Khan. 2019. “Immoral Realism.” Philosophical Studies 176 (4): 897–914.

10. Jonas, Silvia. 2020. "Mathematical and Moral Disagreement." The Philosophical Quarterly 70 (279): 302–27.

11. Jonas, Silvia. 2024. "Mathematical Pluralism and Indispensability." Erkenntnis 89 (7): 2899–923.

12. McPherson, Tristram. 2012. “Ethical Non-Naturalism and the Metaphysics of Supervenience.” In Oxford Studies in Metaethics Volume 7, edited by Russ Shafer-Landau. Oxford University Press.

13. Potter, Michael D. 2004. Set Theory and Its Philosophy: A Critical Introduction. Oxford University Press.

14. Street, Sharon. 2006. “A Darwinian Dilemma for Realist Theories of Value.” Philosophical Studies 127 (1): 109–66.

15. Zalta, Edward N. 2024. "Mathematical Pluralism." Nous 58 (2): 306–32.

16. Zormpalas, Stavros Orfeas. 2025. "How Can the First Order Come First?" Philosophical Studies 182 (11–12): 3085–110.

Schizophrenia and the Neurodivergent Movement

George Getze

Introduction

Many disability rights advocates have been defending an account of disability as not inherently disabling at all, saying that the social structures and situations people with disabilities find themselves in are the root of any disadvantage these people might face. They propose that many disabilities, like learning disorders, autism, blindness, etc., are mere variations in manner of function as opposed to inherently disabling conditions, and that if society were structured differently, these people could find themselves not disadvantaged at all. In this paper, I attempt to argue that a condition that I have, schizophrenia, along with other disorders, are inherently disabling and manifest as essentially detrimental to those who have it. However, I believe we should be able to regard people as disabled in a way that allows us to acknowledge their disadvantage, while also providing support that promotes their function and ability. In addition to this, I will make the point that some disabilities may be considered disabling, or at least significantly detrimental, even if the person with the disorder is high functioning. That is, there is a working notion of disability as hinging on function, but I will try to establish that it can manifest as some other disadvantage besides lack of ability (like profound displeasure). The position one takes on this issue will likely inform their regard for and approach towards neurodivergent/disabled people, so I feel it is crucial that we approach this issue in a realistic way that also takes into account how the people in question wish to be regarded.

Primarily in this paper I will be writing from the perspective of a person who has schizophrenia. My symptoms include paranoia and delusions, with the latter being the more consistent. I frequently believe that I am receiving communications from a higher power, and the content of these messages is generally aggressive and malicious. I believe this 'higher power' is in control of a simulation of sorts in which I am the sole inhabitant, and I often fear I will be subjected to horrific experiences, including my death or worse. This has rendered me disadvantaged relative to neurotypical individuals, and it is primarily from this perspective that I will argue there are many individuals who experience decidedly detrimental mental states yet are high functioning, myself included. Disability (using the traditional definition) is only one aspect of disadvantage which may come with disease; perhaps the more prevalent disadvantage in my situation is just pure discomfort. There are times when I am relatively functional but suffering from frightening delusions which, as a mental state (fear), last for days or even weeks.

In the first section I respond to Anita Silvers, entering into a discussion about how disabled people function given their disabilities. In the second section I engage an article by Mladen Bosnjak and talk about whether we should regard disability as mere difference or disabling, with reference to findings that schizophrenic people in other countries hear friendlier voices than schizophrenics in the United States do. In section three I address another Anita Silvers article, discussing value laden notions of disabled people, proposing that they should be regarded as intrinsically valuable rather than valuable according to their contributions or abilities. I also contribute to a discussion about the quality of life for disabled people, based on my own lived experience. Finally, I address a paper by Richard Dean, responding to the notion that some people may not want medical interventions for their own conditions, or that they may oppose the search for cures and/or further treatment.

I.

In the first article I will look at, *Philosophy & Disability An Overview* by Anita Silvers, there is some discussion of just how disabling disabilities are; specifically there is discussion of how a blind person assesses their situation having been blind since birth, acknowledging a disadvantage but not a complete lack of ability. Silvers says "Milligan (the blind person here) describes himself as an empiricist who does not count being able to see as essential to empirical knowledge, although he readily acknowledges it to be a useful proficiency" (Silvers 2000, p. 38), thus implying that he is disabled relative to his empirical judgements. To address what this issue looks like for schizophrenics, who have a significantly detrimental disability it can become, and at least for me has become, not disabling enough to render us nonfunctional, in cases where medication is effective. I have had the fortune to find a medication regimen that mitigates my symptoms most of the time. I don't mean that I don't have symptoms most of the time, because I do, but they are usually not overwhelming, and I can distinguish them from reality. However, it seems like this is not the case for many schizophrenics. Many people suffering from the disorder do not want to take their meds and without medication, myself included, would likely be nonfunctional in many cases, and would be unhappy and likely suicidal, as I was when I opposed medications post-diagnosis. Schizophrenia is not just a disability that hinders function; it hinders wellbeing and happiness, both because of symptoms and an inability to function. Disorders like this are both difference and handicap; it is the manifestation of an ability most have that some lack, decreasing their ability relative to those who are able (to hear, see, be free of paranoid delusions, etc.)

Silvers also talks in this paper about how disabled people may require additional help to do what other people could do with less help or without help at all. She points out that "allocating the same resources to all individuals does not lead to equal capability" (Silvers 2000, p. 39), raising the notion that different people need different amounts of assistance to achieve the same things. Viewing disability as a reason to exclude the disabled from participation, in what would be empowering activities that they are capable of participating in, with or without reasonable accommodation, is a big part of the problem with calling disabilities disabling rather than mere differences. I believe that we can both empower disabled people and acknowledge that they are at a disadvantage, and perhaps we actually need to acknowledge their disadvantage to accommodate them. If we regard disability as disabling, we can provide the disabled with the extra support needed for them to live better lives at a higher level of function. I have personal experience with accommodations for abnormal circumstances due to my disability. I have been in psychiatric hospitals a few times every year in the past 3 years and was in psychiatric hospitals about 10 times in my first year after diagnosis, 4 years ago. If the schools I have attended hadn't excused my absences and given me a chance to make up missed assignments college wouldn't have been an option for me. Also, there are plenty of other students with disabilities that require more extensive and proactive accommodation than I have needed, who would not be able to get an education at a university without these or other similar accommodations. This doesn't necessitate that people be considered less capable necessarily, especially to the degree that accommodations bring people to a normal level of function, but I think recognizing the disability does not have to harm the person with the disability regardless of whether they function or not. Relative to function, they will likely either be praised or sympathized with as long as the judging person doesn't harbor negative feelings towards the disabled person (as an ableist might). Silvers says "Philosophers like Elizabeth Anderson, Iris Marion Young and myself, who take justice to be about inclusive interpersonal respect and social participation, think the moral challenge of disability is to reshape practices so they no longer exclude people with corporeal or cognitive anomalies" (Silvers 2000, p. 39).

Silvers invokes Eva Feder Kittay, who writes "fear of losing bodily control and fear of suffering may be too deeply embedded in our culture to be dislodged. These cultural taboos prevent disabled people from being honest in communicating their experiences" (Silvers 2000, p. 40). I think we need to construct a narrative around the topic of disability that normalizes both the state of disorder and the expression of the status of a disorder and its symptoms. According to National Institute of Mental Health (NIMH) about 23% of adults in the US have a diagnosed mental disorder. We need to normalize expression of these disorders to open up a community and culture of support for disabled people. Some of them may be independent, and others not, but we all need to be honest and understanding of the barriers we face.

II.

The next paper I will look at is *Is Autism a Mental Disorder According to the Harmful Dysfunction View?* by Mladen Bosnjak, which, like the last paper, does not focus on schizophrenia specifically, but will serve as a baseline for what issues will be addressed relative to schizophrenia. Bosnjak says that the claim that high functioning autism is not a disorder is not supported, because even minor cases of autism involve some disability in social function and mental dysfunction, thus qualifying the condition as a disability in medical terms. I disagree with this notion to some extent and suggest instead that we should assess disability on a case by case basis, using tests of mental capacity to determine disability status (this test may not capture the degree of harm done by disabilities which can hinder an individual without rendering them less capable of completing difficult mental tasks. My disorder is like a hurdle of sorts: it does not affect my own ability to run but it is there and slows me down. This isn't a perfect analogy. Maybe in moments when my schizophrenia is not pronounced I am not really disabled, but when the disorder is pronounced I am disabled). Bosnjak says "I maintain that the list of basic psychological capacities provided by George Graham (2010) provides an adequate conceptualization of harm" (Bosnjak 2022, p. 89), or in other words Bosnjak says that the measure of harm is, in this case, considered to be function or lack of it. I, on the other hand, contend that the measure of disability for schizophrenics should at least in part be set apart from psychological capacities (in terms of capability). I am trying to establish an account of disability that leaves room for significant detriment despite high function. A disorder's impact on affect and mood constitutes a grounds for determining disability status. If a disorder impacts your mood negatively or causes you to be impaired, it is detrimental. This accounts for mood disorders like depressive disorders, anxiety disorders, and maybe bipolar and to some degree schizophrenia. To speak about schizophrenia, if a schizophrenic's delusions are happy and positive, while they may not have the grounding necessary to function in society independently, I think there is something to the notion that they are not being harmed (at least not as clearly harmed) because the delusions are pleasant. This approach seems to evaluate quality of life using utilitarian standards (without the maximally increased value of the higher pleasures suggested by Mill). It is utilitarian in the sense that if something feels good, it makes your life better, and if you feel good most of the time and aren't suffering enough to outweigh the pleasure's impact on the quality of your life, your life is good.

I have also heard that hallucinatory voices heard by schizophrenics are shaped by the societies people live in, being more aggressive in western culture and more playful/gentle in some eastern countries like Africa/India. A study done by Clifton B. Parker, a Stanford anthropologist explores this; an article on this study says "The striking difference was that while many of the African and Indian subjects registered predominantly positive experiences with their voices, not one American did. Rather, the U.S. subjects were more likely to report experiences as violent and hateful – and evidence of a sick condition" (Parker 2014). The medical model of disability focuses on the state of the individual themselves; the social model focuses on how the structure of society puts people at a disadvantage. Although the medical and social models both describe real aspects of hindrance due to disability/social structure, the impact of many if not all psychiatric disorders cannot solely be attributed to social structure or personal state. For Schizophrenics, it appears that cultural and social practices can make delusions more or less menacing, but medical interventions can make symptoms less pronounced. We don't have a clear understanding of why voices for schizophrenics in other countries are friendlier, but even when they are friendly they are still delusions, so I would assume most people would want to rid themselves of delusions regardless of their affective nature.

Bosnjak then introduces the Harmful Dysfunction Account. The Harmful Dysfunction Account says "A condition is a disorder if and only if (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and (b) the condition results from the inability of some internal mechanism to perform its natural function, wherein natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mechanism (the explanatory criterion)" (Wakefield 1992:

384)” (Bosnjak 2022, p. 93). Bosnjak says “it is plausible to think that there are conditions, such as schizophrenia, that are associated with low quality of life, often leading to fatal outcomes, and as such can be considered as harmful regardless of the evaluative standards entrenched in a specific society in which it occurs” (Bosnjak 2022, p. 95). Generally, some people seem to agree that not all disorders are mere differences whose disadvantages are results of the structure of society. I think many of these people would consider schizophrenia sufficiently harmful in a way that is not solely a result of the structure of society to regard it as intrinsically disabling. Maybe what needs to be said is that schizophrenic people, with the right medications, can possibly live relatively normal lives, at least or especially in terms of function. My delusions have never gone away entirely, but when I view my thoughts and delusions through the right lens (one which lends itself to understanding that my delusions are the result of a disorder and are not real) I can have an awareness of what is and isn’t real. That being said, on another level I still believe all my delusions are real; I just tell myself they are not. The real key is that when properly medicated my delusions are less pronounced and less frequent. For some disorders, such as schizophrenia, a person may be high functioning but significantly disadvantaged in areas other than ability. Even if ability isn’t impacted, disorders like schizophrenia can make people less happy and inclined to feel trapped in a life of displeasure, leading to suicidal ideation or suicide attempts. Schizophrenics may well be harmed by stigmatization and biases, but I think probably the most intense and pervasive discomfort comes from the disease itself.

III.

Now I will move into a discussion of another Anita Silvers article, *On the Possibility and Desirability of Constructing a Neutral Conception of Disability*. Given that there is disagreement about where the boundaries of disability lie, I think it should be acknowledged that there are negative and positive stances on either side that people can take, making disability more or less broad. Of course, disabled people may be attributed stigmatized characteristics and be discriminated against; also if disability classifications include more people there may be less stigma and bias carried with it. Disability classifications that reach further can allow more people to get needed help but can also limit the opportunities some people on the fringes of disability have. The cases on the fringes of disability classification (those cases which are close to the threshold separating disability and non-disability) are tricky, and I think people should have the opportunity to keep their diagnoses private. Silvers proposes that we try to develop a notion of disability that is not value or affect laden. This paper raises an important concern: that disabled people might be regarded as lacking in value or unworthy of the extensive support that some of these people need to survive and function. This devaluation can happen at the personal or clinical levels.

Silvers portrays bioethicists as being inherently opposed to disabled people, but I don't think this type of opposition is inherent to bioethics at all. It seems like at least some bioethicists do support care for the wellbeing of disabled people. I think to avoid transgressions against disabled people, including sweeping institutionalization, castration, etc. we do not need to avoid considering disability a bad thing in every case, though for the same reasons I think disability should not be considered inherently bad, or at least not inherently invalidating. The way to stop the propagation of transgressions is to regard people as inherently valuable and not instrumentally valuable. If we do not use people as a means to achieving some other end, we do not castrate them or seek to eliminate them from the population. "In a neutral framework, bioethicists will base their beliefs about the difficulties of living with a disability on facts rather than fears" (Silvers 2003, p. 476).

There is ample discussion of the difficulty in determining whether medical intervention or adaptation to a life of natural disability is preferable in cases of physical disability, and with some mental disorders I can see valid reasons to reject treatment. Silvers makes clear that some people would rather not have treatment for their disorders, "to illustrate, adults with upper limb phocomelia (congenitally anomalous arms)... often believe themselves to have been injured as children by having had their natural digits amputated, been fitted with ineffective artificial arms, and forbidden to use the much more functional method of manipulating objects with their feet" (Silvers 2003, p. 475). However, I don't think all medical interventions should be grouped together for all disabilities. For example, I don't think there is much of anything objectionable in the administration of crutches and wheelchairs, which are medical interventions in a sense. Perhaps bodily invasion is a good standard for measuring how objectionable an intervention is likely to be. Relative to schizophrenia, many schizophrenics don't want to take their medications, but I think many of us who have had the disorder for a while figured out some time ago how much the right medications improve the quality of our lives. There are risks and negative side effects that come with taking antipsychotics, but I think in the case of those drugs, others, and some other medical interventions, there is nothing obviously objectionable about them. Of course, though, there are some interventions that are potentially objectionable, such as those administered at birth when the patient is unable to give any consent. Silvers makes the point that not all disabilities are health issues: e.g. learning disabilities like dyslexia and ADHD. Silvers also says "Disability advocates who call for giving the disabled priority in the distribution of resources in virtue of being definitively the worst-off members of society likewise may fear adopting a neutral conceptualization of disability" (Silvers 2003, p. 477). I think the notion that disabled people are definitively the worse-off members of society is probably a contestable claim, the status of which depends on how much gravity different aspects of one's life warrants. For instance, plenty of people are more openly opposed and subject to prejudice in society, e.g. members of LGBTQ society, people in minority racial and social groups around the world, etc. While disabled people struggle with the hindrances of their disability, most people do not hold a disabled person's disability against them overtly. Consequently, the negativity they experience is not rooted in any hatred of disabled people (at least not often openly in America). Silvers says that people should not get priority treatment due to disability, but due to physical or medical need. People can have disorders that are not relevant to the treatment they could receive preferentially because they're disabled, e.g. a person in a wheelchair skips waiting in line. This is not necessary because the wheelchair may be comfortable, but if a person came in who was walking but with difficulty, they should be able to skip the line.

Silvers says "the source of the badness of disability commonly is thought to be its connection with physical, sensory, or cognitive impairments" (Silvers 2003, p. 479). I would like to bring to consideration disorders that bring about new, detrimental symptoms without impairing any prior capability. As my disease progresses, I find, more and more, that this has become the circumstance I am in with schizophrenia; that is, my senses are not impaired, but I experience hallucinatory or delusional positive symptoms that are experienced as perceived or thought of. Four years ago, when I was first diagnosed, I was significantly cognitively impaired. However, I regained the ability to function cognitively in relatively short order. However, I was still experiencing the delusions and hallucinations that are most characteristic of my disorder. The point is, my function might not be impaired now and may not have been impaired much for the last 3 years or so, but the positive symptoms I am still having (delusions/hallucinations) make my life significantly harder. I think it is most accurate to regard many disabilities, including my own, as decidedly detrimental. However, thinking that way about my condition (which I don't often do) makes me feel sad, helpless, and pitiful. I think that we need to acknowledge the disadvantages disabilities impose on people when trying to figure out how to help them make their lives better when they are struggling, without having it bring them down emotionally when they hit their stride. I should logically be able to countenance that I am doing very well given my circumstances, in my personal, professional, and academic life, but the reminder that my disease tends to dramatically reduce quality of life for many patients makes me wonder what my life would be like if I wasn't schizophrenic and makes me miss my life before schizophrenia. In any case, to raise people up from the depths of, in this case, schizophrenic delusion we need to provide support (and a good medication regimen). For high functioning disabled people, maybe it is more harmful to remind them of their disabilities and how negative the consequences of having their disability can be; maybe we should simply provide support.

Silvers presents a notion to go against the thought that just because one alternative is preferable to another the less preferable is inherently bad: say there is someone richer than you, they have 2 billion dollars to your 1 billion. Does this mean that having just 1 billion dollars is objectively bad? Clearly not, but this thought made me think maybe there are two possible ways of determining general 'goodness' or 'badness' of a life. One is to just compare the quality of one's life to the average quality of life for members of a species, every living thing out there, or for some baseline group. However, I think the better measure of quality of life is a self-determined life quality. It is possible that the least happy human on the planet could still be happy at some point in time or even all the time (it is at least conceivable that this could be the case in some world/scenario). So given this, we should listen to what disabled people say about the quality of their lives while taking into account how long a time they have had to adapt to their impairment/dysfunction. The notion of the negativity of disability fuels their marginalization and oppression by dominant groups (as is the case for so many other oppressed groups), but this needn't be the case. We can acknowledge a lack of ability as a reason to assist and uplift disabled people instead of a reason to marginalize them or eliminate them from society or keep them hidden away in institutions. Many marginalized groups are considered inferior, and disabled people are in many cases less capable than able bodied or minded people, but they are not strictly oppressed because of the notion of their inability; change needs to come in our methods of addressing the disadvantages we find people at.

Silvers points out that we have largely constructed a neutral notion of women's differences, which considers their differences neither positive nor negative. However, I feel like with some disorders (such as my own) the neutral notion cheapens the significance of the disabilities. I will say, the neutral conception of disability can be empowering, and maybe when paired with proper treatment (without stigma accompanying the need for treatment) can bring disabled people closer to a normal, equitable life. This notion is strengthened by the discussion of the neutral conception of women, however I think the success of this notion is largely powered by women's ability to be independent and just as functional as men in the areas of life where function most matters (in work, school, family life, etc.). If we simply consider disabled people's differences neutrally, we might be shifting a burden of expected decreases in dependence on the disabled people we are treating more equally (we might be pressuring them to accomplish an independence they may be incapable of sustaining). I think the inability and vulnerability of disabled people is crucially important to recognize as the cause and justification of their increased dependence, while healthy women are much more capable of being independent and of being as successful as men, warranting neutral regard for their differences. Silvers also asks, "in what respects is society obligated to accommodate individual differences?" (Silvers 2003, p. 485). Society is not necessarily obligated to accommodate everyone or anyone at all, however it is morally just and in the best interests of society as a whole to provide actual support and produce a sense of support and continued opportunity for people in difficult situations caused by disability. This can be connected to a Rawlsian approach to structuring society: we should make rules we feel are just without knowing who we will be in a society, such that, hopefully, no matter who we are we would feel like we are treated justly in that society. This translates to a realistic situation where many of us may find ourselves; a lot of disorders develop later in life, so we may live some time at a distance from disability only to find ourselves a part of some group of disabled people after the onset of some disorder. This is where a sense of security comes into play: we should live with some peace of mind knowing that we will be supported by society even if we come down with a disorder.

IV.

The final article I will look at is *Neurodiversity and the Rejection of Cures* by Richard Dean. Dean talks about how some neurodivergence and disability advocates reject personal treatment and even in some cases the search for better treatment and cures generally. He doesn't take a hard position on all the issues addressed, but assesses each position individually. I believe, based on personal experience, that opposing the search for better treatment or the search for cures to schizophrenia is ridiculous and wrong for all the people who could lose out on treatment that could drastically improve the quality of their lives. That is to say, I think most schizophrenics would live better lives if they accepted treatment, though they have a right to not do so. But opposing the search for cures and better medical treatment is wrong because it robs those who want treatment of more effective medicines or cures. I have firsthand experience living with schizophrenia while rejecting medical intervention at one point and then accepting it later. It has been my experience that medical interventions have made my schizophrenia bearable and have made me feel like life is worth living. For me, living with untreated schizophrenia consistently led to suicidal ideation. Some people may reject medications, and if they feel their life is worth living without meds then they have the option to not take medication. There are a lot of schizophrenics who struggle to find life worth living because of their symptoms. It would be cruel to disregard these people's dependence on their medications or to stop searching for cures or to stop giving them desperately needed and highly valued/desired treatment.

Dean suggests that an objection to medical intervention is that with conditions like autism, much of a person's identity is intertwined with their condition. The notion is that autistic people are not, in a manner of speaking, trapped in a disabled state from which their identity can be removed; their autism is intrinsically a part of who they are. However, I believe this is less obviously the case for people with disabilities that develop later in life, since there was at least at one point a complete version of a, say, schizophrenic person who existed before their condition developed. Their disability can be regarded as an integral part of who they are after it develops, but they can possibly conceive of life without it. I think many schizophrenics are likely to prefer their lives prior to the development of their condition. This would be less like taking away an intrinsic part of their identity and more like a return to a previous state, such that their disability is something they are trapped in and which can be removed without them losing an integral part of their identity. Like with autism, schizophrenia shapes and colors much of what a person with the disorder thinks and controls how they process information, but I at least, as a person with schizophrenia, think of those thoughts as the result of a disease that I don't want to or actively consider a deeply seated part of my identity. Maybe it is an integral part of who I am, but I don't actively consider it a crucial part of who I am or want to consider it a crucial part of who I am.

Dean also points out that many of the people at the center of the neurodiversity movement are very high functioning. This raises a question that I feel should be of utmost importance, since this is what I feel the dialogue surrounding the neurodiversity movement should be dealing with: what is the best way to approach the situations that significantly disadvantaged disabled people are in? Sorry for being blunt, but I'm not really concerned about what is best for high functioning, independent, low severity autistic people when there are people much worse off who are feeling the consequences of their advocacy. I don't think regarding disability as mere difference is inherently damaging for severely disabled people, but I also think it has the potential to be. I think our way of approaching disability should first and foremost help those that need it rather than those who don't need it. From there, we can find ways to treat high functioning disabled people the way they want to be treated; they can be regarded as inhabiting deviations from the normal state people with any particular disorder may be in. But again, the point is simple: we need to make sure people who need the most help get it before those who need it the least.

Dean also says that seeking a cure for autism may send a message of disrespect and devaluation to people with the condition. I think this is a valid concern, but also that it is not a necessary part of the thought driving a desire to find treatment or cures for disorders in general. One can foster complete respect for people with a disability and simply seek out treatment advances and cures out of concern from their wellbeing. I think what is crucial to recognize is that the respect or disrespect people feel towards people with disabilities is generally established prior to a decision to seek out cures or not. It is reasonable that people with disorders might regard research efforts as sending a message of disrespect inherently, but I think this concern can be settled if we as a community express our acceptance of and support for people with conditions we are seeking cures for. The key is that, even if seeking a cure might seem implicitly disrespectful, we as a society can show support and dispel any concerns about disrespect or devaluation.

Conclusion

Despite the argument I have been constructing here, there are compelling reasons to treat disabilities as mere difference. To think of oneself as disabled and less capable as a result is not a happy thought to have and is disempowering. Thinking of disability as mere difference opens the door to a more equitable notion of oneself and notion of others who are disabled. However, a failure to countenance the real disadvantages that people with disabilities face could deprive them of needed treatment and assistance. It is a tough issue to address, and I think the way we decide to address needs to prioritize the wellbeing of disabled and nondisabled people alike, with a relative priority granted to the wellbeing of disabled people, given they are the ones in need of assistance and the ones subject to the manner of treatment we are trying to establish as appropriate. I think we may need to view their condition as mere difference sometimes, and as putting them at a disadvantage at other times. This split in manner of treatment should service the goal of providing opportunity and respect for neurodivergent/disabled people.

Bibliography

Bosnjak, Mladen. (2022) “Is Autism a Mental Disorder According to the Harmful Dysfunction View,” *Croatian Journal of Philosophy* 67, pp.89-111

Dean, Richard. (2018) “Neurodiversity and the Rejection of Cures,” *Disability in Practice: Attitudes, Policies, and Relationships*, pp.115-133

National Institute of Mental Health. (2024) “Mental Illness,” *National Institute of Mental Health,* *www.nimh.nih.gov/health/statistics/mental-illness*.

Parker, Clifton B. (2014). Hallucinatory ‘voices’ shaped by local culture, Stanford anthropologist says. Retrieved December 30, 2025, from Stanford Report: https://news.stanford.edu/stories/2014/07/voices-culture-luhrmann-071614

Silvers, Anita. (2000) “Philosophy and Disability An Overview,” *Philosophy Now, pp.38-41*

Silvers, Anita. (2003) “On the Possibility and Desirability of Constructing a Neutral Conception of Disability,” *Theoretical Medicine and Bioethics 24, pp.471-487*

ROUGH IS REAL: TRANS IDENTITY DOES NOT HINGE ON FIRST PERSON AUTHORITY

Han Kirkegaard

INTRODUCTION

One foundational goal of trans philosophy is to enable the whole spectrum of trans[1] people to find authentic expression for their transness.[2] A popular philosophical vessel for following this lodestar is first-person authority (FPA), the thesis that avowals arising from introspective assessments about internal states are authoritative. As such, arguing for justifications of FPA to anchor trans identity claims is a popular area of interest in trans philosophy. However, in this paper, I will argue that seeking justifications of FPA over gender identity claims is a philosophical trap that obscures the existent practices that do make trans identities real. I begin in Section I by looking at two plausible justifications of FPA: ethical FPA as articulated by Talia Mae Bettcher (2009) and epistemological FPA as argued for by Gus Turyn (2023). I make clear the shortcomings of each account to motivate my move in Section II to bring gender identity avowals "back to the rough ground" as Ludwig Wittgenstein advocates in *Philosophical Investigations* (2009, §107). Having reframed avowals, I claim in Section III that Wittgenstein's private language argument has explanatory power for how it is that trans identity claims are believed without requiring any extraordinary justification for FPA. To this end, I draw an analogy between gender identity and Wittgenstein's beetle in a box. Furthermore, I contend that this descriptive account better tracks with the lived realities of trans people. In Section IV, I defend my account against a relativist objection and in Section V, I return to Bettcher's ethical considerations to remedy the lack of political potency in my account.

SECTION I: JUSTIFICATIONS FOR FIRST-PERSON AUTHORITY

FPA, in a general sense, says that we should believe people when they tell us things about how they experience themselves, who they believe they are, and what they think and feel. These sorts of personal avowals are, importantly, thought to be inaccessible from any but the first-person point of view. That is, FPA reigns over the private inner world of the self, enforcing others' belief of one's first-person claims on the basis that, without access to other minds, there are no evidential grounds to evaluate these expressions of otherwise private states. The presumption of the truth of first-person avowals is traditionally rooted in a Cartesian view of the mind, where the mind has perfect access to its own contents. However, as will be discussed, alternative approaches have been constructed to justify FPA's imperative to believe.

Trans identity claims, taken as an expression of an otherwise private gender identity, fall nicely within the domain of FPA.[3] As outlined above, there are good reasons to seek a normative force for believing trans identity claims. To start, it is hard to imagine what, other than FPA, might be up to the task of bringing the diversity of trans experiences under one affirmative theory. Furthermore, what appears to be at stake is dire—to doubt the authority of trans people to stake claim to their gender identity is to fundamentally violate the autonomy of trans people. This is the core of Bettcher's (2009) thorough and compelling argument for FPA being a matter of ethical obligation.

Bettcher begins by casting doubt on the possibility of an epistemic justification for FPA. She presents the example of a psychoanalyst correctly determining from knowledge of a patient's behavior that the patient's avowals of love for and trust in her brother are not reflective of her actual feeling—one of betrayal (Bettcher 2009, p 100). In this example, first-person knowledge of how one feels appears deficient relative to a third-person observation of a supposedly private internal state. This serves to show that there is sufficient reason to doubt that we have the kind of unimpeachable access to our thoughts, feelings, beliefs, and desires that an epistemic justification for FPA would require.

Bringing the example closer to the topic at hand, I would make the further argument that the dominant narrative of trans experience depicts a period of deep confusion about one's own gender prior to the realization that one is trans. It is hard to see how privileged epistemic access to one's own gender identity could be so painfully absent for a period and then unquestionably authoritative afterwards. It could be argued that the initial confusion is the product of cultural norms that oppress the trans individual's innate gendered self, hiding it from the mind's clear view. In Section III, I will return to the problems presented by any argument that relies on the innateness of a gendered self. For now, it is enough that trans people don't, until they somehow do, have authoritative epistemic access to their gender.[4] Consequently, for FPA over gender identity claims to stay afloat, there must be an account of how epistemic access can be clouded and then restored.

Wary of the now compounding work needed to shore up epistemic FPA, I follow Bettcher in concluding that an epistemic justification for FPA is too shaky a philosophical position to hold.

With epistemic FPA discarded, Bettcher moves on to a more substantial base to hold up FPA over trans identity claims: ethical FPA. The argument for ethical FPA proceeds as follows.[5] We have an ethical obligation to protect a person's autonomy (Bettcher 2009, p. 103). Autonomy is compromised when one can only choose one way or else be punished. Gender systems that punish those who present a gender not aligned with their genital status do not allow agents autonomy over their gender presentation. Furthermore, in the case where gender presentation is a system of genital representation, enforced compliance amounts to forced disclosure of genital status (Bettcher 2009, 106). We have strong moral taboos against nonconsensual genital exposure (Bettcher 2009, 107). Therefore, reality enforcement of gender presentation not only restrains the autonomy of the person, it also transgresses her moral boundaries (Bettcher 2009, p. 115). As such, we have an ethical obligation to resist reality enforcement of gender presentation by treating a person's avowed gender identity as real (ibid.).

Here, Turyn (2023) picks up the argument. While the ethical dimension of FPA Bettcher puts forth is strong, it appears to stop at enforcing behaviors with nothing to say about belief. Turyn says, "when we merely act as if we believe others about their gender identities despite not believing them, we seem to be doing something wrong, and this does not seem to be accounted for by the dominant understanding of FPA" (Turyn 2023, p. 3). It is from this sense of wrongness that Turyn argues an epistemic norm of FPA is needed (ibid.). Importantly, Turyn is operating from the premise that avowals are the crux of credibility for gender identity, though it is notable that this is only applied to trans identities.[6]

I take Turyn to be correct in teasing out a sense that something feels not quite right about ethically unimpeachable yet insincere actions. However, Turyn's examples only seem to plausibly demonstrate that acceptable behavior masks conflicting belief in low stakes social interactions. See Dana and Chris: Dana choses to use the correct pronouns for Chris (nonbinary) to not offend them but does not believe in gender outside a cisgender binary (Turyn 2023, p. 8). Turyn wants to say that Dana should believe Chris because Chris's avowal is the best possible evidence for their nonbinary identity (Turyn 2023, p. 16). If belief can be enforced from an epistemic argument such as this, then the gap between belief and action would be resolved. But, on a closer inspection, it is unclear what belief would mean in this case. How is Dana supposed to understand what it means to be nonbinary even if she were to accept that Chris's avowal is decisive? If she doesn't understand what it means for Chris to be nonbinary, what is there for her to believe?

Imagine that Chris is the first trans person Dana has ever met. She doesn't believe in gender outside of a cisgender binary when she meets Chris because her experience has limited her to that belief. Over time, as Dana and Chris become friends, Dana has to think less about using the correct pronouns for Chris. She becomes familiar with the soft way Chris speaks, she notices the way they sit with their hands folded in their lap, she unselfconsciously shares her feelings with Chris and feels seen in the way Chris carefully responds to her. Dana hardly remembers when they first met and Chris had to correct her on the pronouns she used for them. It is less that Dana believes Chris is nonbinary, she just knows Chris. One day, an acquaintance, Sam, misgenders Chris to Dana. Dana corrects Sam only for Sam to respond, "But you know what I mean. Chris is really a ___."[7] Dana thinks about it. This far into knowing Chris, she can no more convince herself that Chris is really a ___ than she could believe Chris was nonbinary when they first met. She knows what Sam means, that Chris has genitalia associated with ___. But even still, she can't think of Chris as a ___ because she doesn't experience Chris as a ___ or treat Chris as a ___.

This counters Turyn's approach in two ways. First, it hardly seems that Chris's avowal of their gender identity is what caused the change for Dana. Second, it seems that the issue does not stem from a distinction between action and belief, since the two collapse into one. Through the intimacy of friendship, accumulated over time, Dana has become so familiar with the behaviors of Chris that Chris's identity has become real for Dana. Following this, an epistemic norm for FPA is not necessary to solve the problem. Dana does not need more confidence that Chris's avowal dutifully represents some internal gendered state, she needs more confidence that she knows what Chris means when they avow their nonbinary identity. By knowing Chris—their mannerisms, their temperament, their values, and so on[8]—Dana does come to know what Chris means when they say they are nonbinary.

Moreover, Dana no longer thinks she should believe Chris is nonbinary, she just sees them and treats them as nonbinary, having come to grasp that gender presentation through her time around Chris.[9] Returning to Bettcher's (2009) ethical FPA, this account of Dana and Chris gives no reason to doubt that we have an ethical obligation to act in accordance with FPA over gender identity claims. If anything, it supports it. Dana would have been unlikely to develop the kind of relationship with Chris that made Chris's gender real to her if not for her choice to treat Chris in accordance with their avowed gender identity. However, Bettcher's account doesn't explain why Dana stops feeling she *should* believe and how she instead comes to be unable to *disbelieve*. Thus, where Turyn has gone wrong is in hinging credibility on avowals alone and what Bettcher is missing is a richer account of how avowals are part of an evidential matrix of behavior and how this dictates belief in a non-normative sense.

SECTION II: BACK TO THE ROUGH GROUND OF AVOWALS

This section takes seriously the following view on language from Wittgenstein (2009):

> The more narrowly we examine actual language, the sharper becomes the conflict between it and our requirement… We have got on to slippery ice where there is no friction and so in a certain sense the conditions are ideal, but also, just because of that, we are unable to walk. We want to walk: so we need friction. Back to the rough ground! (Wittgenstein 2009, §107)

What follows is an attempt to retrieve avowals of trans identity from lofty philosophical discourse and bring them back in contact with the practices that generate and sustain them. With a realistically textured view of avowals on hand, the need for a norm of FPA will be significantly deflated. The resulting theoretical underpinnings of gender will be supported in the subsequent section.

A broad, informal survey of the chaotic landscape of how people talk about trans people reveals a surprising consensus—anti-trans jokes through trans-authored narratives seem to agree that self-avowals about gender identity are at least consistent with what a trans person is observed to be like and perhaps even essential to what it is for someone to be trans. Think of the transphobic joke: "I identify as an attack helicopter." Now look to the stories trans people tell about themselves. Coming out narratives abound. The hazy idea that saying who we are is important, or even necessary, to who we are has solidified in the uneven ground of ordinary trans discourse. Notably, this remains hazy elsewhere. Gender identity avowals are rarely if ever part of strictly heteronormative discourse due to the naturalization of the stacked gender on sex binary. That the significance of self-avowal is simultaneously such a pervasive and distinctive idea about trans people warrants close attention. However, as I will argue in this section, that does not necessitate that avowals alone entail the necessary and sufficient criterion by which gender identity claims are validated.

Given that avowals are everywhere in trans life, what do they do? The coming out narrative fits closest with how avowals are taken philosophically, that is, as an epistemically private report in need of justification (epistemic FPA) or as the staking of a claim that one can be held responsible for (ethical FPA). As the story tends to go, after some period of reflection, Chris would achieve clarity and resolve regarding their gender identity and would then report this achievement by coming out. This clarity, which precedes the coming out avowal, is what epistemic FPA appeals to. Consequently, interpreting this clarity as unimpeachable self-knowledge, belief in Chris's avowal would be justified. Alternatively, interpreted through an ethical FPA framework, Chris's coming out would be an expression of accountability for their gender identity in opposition to a coercive system of genital representation such that to not believe Chris would be to deny their autonomy and violate their privacy. However, both accounts leave open the question of what it is for Chris to know their gender identity or take accountability for it.

Trans subcultures offer insight into filling this gap and reframe the significance of coming out. In the case that Chris's friends are familiar with trans identity language and practice, Chris's coming out would not convey new meaning in the ways described above. Chris coming out as nonbinary would likely only formalize what is already known. This is captured in the trans slang term "egg" used to refer to trans people who haven't yet realized they are trans. Chris's behavior might prompt the usage of "egg" if Chris's behavior is seen as consistent with expectations for being nonbinary before they commit to that identity themself. In this schematic, coming out would only grant permission for Chris's friends to change the pronouns they use, but would not reshape the way they see Chris. What this shows is that avowal is not separable from the behaviors that precede and follow it. The process of acquiring a position from which one is ready to make the avowal will, in some cases, already have done the work of deciding that Chris is nonbinary upon their avowal. In that way, an avowal cannot entirely reshape how a person is seen just as no one can entirely change the meaning of a word with a single utterance.[10]

Gender identity avowals extend beyond the initial coming out story. Another application is a resistant, continuing mode of avowing a gender in defiance of the naturalized sorting function of the gender binary. When Chris and Dana first meet, Chris corrects Dana's use of pronouns for them. Dana has assumed a set of pronouns that refer to Chris, and Chris has avowed their gender to combat this assumption. This case illustrates that trans people avow their gender identity because it is not naturalized in language. This resistant deployment of avowal explains why the phenomenon does not cross into cis-heteronormative discourse. Trans people have to use avowals otherwise their gender will be assumed in dominant cis-heteronormative discourse, often incorrectly. Cis people don't have to use avowals because their gender will be assumed in dominant cis-heteronormative discourse, often correctly.

Here, I take it as necessary to consider a principle of parity. By a principle of parity, I mean that philosophical accounts of trans existence should require that any explanation or justification of a feature of trans experience be fairly matched with a plausible, analogous cis case. Thus, any account that relies upon some special kind of self-knowledge or certainty in a trans case that isn't plausible in a cis case violates this principle and is suspect. Trans people are not a different kind from cis people. Both are people. They differ with respect to their relationship to sex and gender.[11] It is not assumed that when a cis person doesn't avow their gender identity, that is because they are uncertain about their gender or aren't taking ownership of their gender. Cis identity is stable with or without avowal in this context because it is established through behaviors and embedded in the practice of language. So, when a trans person resistantly avows their gender, they can't be taken to be expressing metaphysically demanding certainty about their gender. They can only be offering behavioral evidence of their gender that is on par with the implicit behavioral evidence embedded in cisgender practices. Thus, what is shown from trans identity claims used in this way is only that the sex and gender binary is naturalized in language, thus prompting the need to intervene.

Here again, Bettcher's account comes out right: avowals of gender identity convey information about how to behave but don't on their own seem to function to demand belief. If this assessment of avowals is correct, then there is good reason to believe that gender is not the sort of thing that has meaning because it is believed but rather that it is something that is believed only when it has meaning.

Section III: Gender Cannot Be the Beetle in the Box

Through an analogy with Wittgenstein's beetle in a box, this section explains how it is that gender avowals can be meaningful and thus believed. Wittgenstein sets up the thought experiment: "Suppose everyone had a box with something in it: we call it a 'beetle'. No one can look into anyone else's box, and everyone says he knows what a beetle is only by looking at his beetle" (Wittgenstein 2009, §293). The beetle is characterized only by being epistemically private, but one could imagine that each person has perfect epistemic access to their beetle. This bears a strong resemblance to the metaphysical sort of gender presumed by an epistemic conception of FPA. So, imagine this beetle is the kernel of gender that, in an epistemic conception of FPA, would endow gender avowals with credibility. Like the beetle, no one can show this kernel of gender to anyone else, everyone is alone with their innate experience of it. Wittgenstein goes on to say that there would be no way for 'beetle' to have a use in language, "not even as a *something*" (ibid.). Likewise, if gender really referred to an innate internal experience, it could not have a place in language.

Such a strange claim, or at least one that runs against so much of traditional philosophical thought, arises from Wittgenstein's rejection of ostensive definition and his claim that meaning is "not only agreement in definitions but also (odd as it may sound) in judgments" (Wittgenstein 2009, §242) where such agreement is "agreement not in opinions, but rather in form of life" (Wittgenstein 2009, §241). For Wittgenstein, the very act of speaking a language constitutes a form of life (Wittgenstein 2009, §23). What this says is that, for language to have meaning, it must have shared criterion for adjudication. Without this, there would be no way of teaching or demonstrating mastery over the use of language for there would be no way of demonstrating that the language was used correctly.

Consider, it may be an empirical matter of fact that trans people spend more time than cis people thinking about their gender and therefore would be more familiar with the contours of their internal experience of gender. If this does come out true, it would then seem plausible that this would render a richer inner phenomenal experience of having a gender. If you are inclined to call this robust acquaintance with subjective, interior gendered experience a kind of self-knowledge, then it follows that trans people are well positioned to have more self-knowledge about their gender than cis people do. Following this, trans gender identities should be believed with more confidence than cis gender identities. However, as discussed above, cis identities are naturalized in language in a such a way that cis gender identities always appear less in need of justification, and thus are more readily believed. This shows that our language is ill-suited to accommodate this proposed trans epistemic advantage. If language really did function as a report about inner states, then language should articulate something about self-knowledge and such an argument for confidence in trans identity avowals would be apparent in ordinary discourse. Given the clear evidence that trans identity claims are almost never believed with more confidence than cis identities, it follows that gender, as it arises from our language, is not about a private internal state.

Returning to the beetle in the box, the conclusion is consistent. The degree to which one knows their gender, in the mode of self-knowledge outlined above, cannot be the criterion by which one can be said to have that gender. Even perfect epistemic access to one's private beetle does not let one say anything about how one's beetle could be similar to or different from anyone else's beetle. For gender to be like the beetle would mean that gender would be totally unintelligible. There would be no way to know that anyone shared in this or that gender. There would be only a proliferation of genders. Every one a private beetle in a box.[12]

But we do meaningfully talk about gender in our ordinary language. That there is disagreement over what counts as correct usage of gender identity avowals is evidence that there is in large part agreement in meaning. A dispute would be impossible if the grounds for that dispute weren't already established. This bedrock of intelligibility is the sharing of a form of life. Thus, trans avowals of gender identity don't express a private reality and in fact lie fully in the domain of intelligible gender discourse. This does not deny that one can have subjective experiences related to gender. Only that, insofar as these feelings attach to a communicable meaning, there is a public facing expression of them. This public facing expression is how language traffics meaning in.

Section IV: Language is Regular

One might object that this leaves us in a place where we can't really speak about particular gender categories at all. If there is no metaphysical anchor or unchanging referent for the gender categories we use, then that untethered meaning will float away into something immaterial and counterproductive. The way this worry is actualized is in the complaint that feminism dies by no longer having a definite class of subjects to liberate. But this presupposes that what is operational in our language is so because the words pick out some distinct set of necessary and sufficient conditions to render it meaningful. This is far from proven and in fact, the opposite has been demonstrated above. Given how much of our language gestures at things that cannot be given a satisfactorily concrete definition, it has to be accepted that language can—and does—function without picking out clear boundaries around the meanings of words. The basis for meaning then is intersubjective agreement. Intersubjective agreement demands regularity or else there could be no agreement. Trans identities do not undermine this, they operate resistantly to the naturalization of binary sex determining binary gender but that is still while conceding many of the more basic ways in which gender functions. For example, when a trans woman is seen as a woman, she is seen that way because of the pre-established practices of being a woman. She no more reinvents womanhood than any cis woman does.

This regularity stipulation for gender identity avowals appears tacitly in the background of Bettcher's (2009) argument and is ultimately compatible with her conclusion. She writes of trans subcultures: "self-identifications are generally accepted at face value" and this is the case even though "identity terms (*transman, genderqueer*) do not have well-specified, fixed definitions" (Bettcher 2009, p. 108). This regularity is guaranteed because agreement in language is negotiated. It happens incrementally over time and not without dispute. But it isn't arbitrary. As Wittgenstein says, "One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'" (Wittgenstein 2009, §258). "Right" like meaning is a matter of agreement. Women will not cease to mean something, even if this something has elements of dispute, unless gender ceases to have use in our language, which is unthinkable from our vantage point.

Wittgenstein, considering the complaint that language can't have such "blurred" edges offers this: "Is it even always an advantage to replace an indistinct picture by a sharp one? Isn't the indistinct one often exactly what we need?" (Wittgenstein 2009, §71). If gender categories had perfectly resolved boundaries, this would either make no one a real woman, or else for anyone to be a woman would be for that woman to express their gender in one universal way. Blurriness is generative. It engenders degrees of freedom in systems of categorization. This allows for interpretation, creativity, and change. To be sure, it also does, at times, lead to confusion and mistakes. Yet it functions.

SECTION V: RETURNING TO ETHICS AND THE POSSIBILITY OF POLITICAL POTENCY

As a descriptive account, following the methodology of the later Wittgenstein, this account does not have a normative force. It merely attempts to clarify what already is. However, given the harmful political climate for trans people, it is worth considering what would be required for this account to be politically potent. To this end, I return to Bettcher's (2009) ethical FPA with a sharpened focus. Given that belief doesn't appear to be governable in accordance with a norm and a behavioral norm is, as Turyn (2023) points out, insufficient. I propose a complementary ethical framework to supplement ethical FPA. Ethical FPA says: one should treat trans people as if their gender identity avowals are authoritative. This complementary norm, which could be called ethical first-person doubt (FPD) says: one should doubt one's own authority in adjudicating another's trans identity claims. Everyone, even those embedded in trans subcultures harbor potentially harmful biases and patches of myopia. The hostile political climate is such that many trans people do not have access to the material conditions necessary for them to assume some of the behavioral roles we anticipate from trans identity claims. This requires a more in-depth discussion and defense but at least presents one potential path towards an affirmative trans politics that is neither metaphysically burdened nor neutral on the issue of belief.

CONCLUSION

FPA over trans identity avowals, despite appearing promising for trans philosophy, is untenable under a Wittgensteinian conception of language. This is because gender identity avowals are not a special kind of report about a private mental state, as epistemic FPA would require. Moreover, while ethical FPA is unobjectionable as a foundational part of a normative system, it alone cannot explain how avowals of gender identities are already made real, albeit only in the small subset of ordinary practice where trans subcultures reside. Additionally, ethical FPA remains silent on belief. To remedy this, I have presented a descriptive account that draws out the way in which avowals function as part of an evidential matrix that determines belief. This strips the authority from FPA over gender identity claims and repositions gender as part of a shared and publicly adjudicated practice with no clear category boundaries.

End Notes:

[1] Throughout this paper, I use the term trans to capture any gender identity that falls out of the strictly cisgendered binary. This includes, non-exhaustively, binary trans, nonbinary, two-spirit, intersex, and genderqueer identities.

[2] See Stone (1992). Bettcher's (2009) chapter opens with the same reference. Stone's role in shaping the development of trans studies and trans philosophy is just that strong.

[3] This view of trans identity claims will be dismantled through the course of this paper. However, it is useful to start from this view of trans identity claims to get a sense of what FPA is and what it is meant to do.

[4] Some trans people do feel as though they have always known their gender and may not agree that transitioning to that gender identity ever involved any confusion of the sort described here. However, returning to Stone's (1992) summoning of the multiplicity of trans voices, employing this kind of certainty as a justification for an epistemic claim to FPA risks invalidating any trans person who fails to meet that standard. Therefore, this only goes further in demonstrating that epistemic FPA will not be a satisfactory tool for trans philosophy.

[5] This glosses over some of the nuance in Bettcher's (2009) argument, such as the role of responsibility for avowals and the different senses in which privacy structures moral boundaries. While these details add richness to the argument and, in fact, have been generative for the development of my own argument, I have stuck with a simpler argumentative structure here for the sake of clarity.

[6] More on this asymmetry with cisgender identities in Section III.

[7] Here, "___" functions as a placeholder where either male / cis man or female / cis woman could be inserted.

[8] This is the "who" rather than the "what" of Chris's identity, to draw from Bettcher's "existential identity" (Bettcher 2009, p. 99).

[9] There are many alternatives to how Dana might come to see Chris. She might see Chris as really an exception to Chris's presumed "natural" gender category or as really gay but not nonbinary or as really just radically confused. Nothing in my account necessitates that Dana see Chris in

alignment with their avowal, though I find it hard to imagine a genuine friendship existing between the two if Dana does not at least approach sincere belief in Chris's gender identity.

[10] This is both intuitive and a consequence of rule following in language games as described by Wittgenstein (2009), see §80 through §85 (p. 38-39).

[11] In this paper, I leave this principle of parity mostly undefended in part because I take it to be convincing even as it is stated here. That said, it seems worth developing further.

[12] It is funny that in endorsing the innateness of gender, the anti-trans advocate inadvertently conjures her worst nightmare: a world overstuffed with subjective gender experience with no intersubjective agreement about what any of them could have in common. A world of "I am because I say I am."

Bibliography

Bettcher, Talia Mae. (2009) "Trans Identities and First-Person Authority," in: Laurie J. Shrage, *You've Changed: Sex Reassignment and Personal Identity*, pp. 98 – 120 (New York: Oxford University Press)

Bettcher, Talia Mae. (2024) *Beyond Personhood: An Essay in Trans Philosophy* (Minneapolis: University of Minnesota Press)

Stone, Sandy. (1992) "The 'Empire' Strikes Back: A Posttranssexual Manifesto", *Camera Obscura: A Journal of Feminism, Culture, and Media Studies* 10(2), pp. 150-176

Turyn, Gus. (2023) "Gender and First-Person Authority." *Synthese* 201(4), p. 122.

Wittgenstein, Ludwig. (2009) *Philosophical Investigations, 4th edition, trans. Hacker and Schulte* (Wiley-Blackwell)

INDETERMINACY, SUPERVALUATIONISM, AND POSSIBLE WORLDS

Diego Irigoyen Monteiro

INTRODUCTION

Philosophical matters of vagueness overlap with paradigmatic issues in metaphysics, language, and logic (such as composition, mereology, excluded middle, and others[1]). The literature splits the possible sources of vagueness[2] into three types: semantic, epistemic, or metaphysical, i.e., worldly vagueness. Seemingly, the first of these types represents the most commonly held position, while the second and third are considered minority views. Of these minority views, metaphysical, ontic, objective, or worldly[3] vagueness is the view which holds that the source of (some or *all*) vagueness is the non-representational world, or, put differently, the world itself. Within the various versions of this worldly view, Barnes and Williams (2011) and Akiba (2000 and 2004) propose treatments of vagueness as a sort of modality extending over a range of possible worlds. These philosophers also coincide in the presentation of their views as both alternatives to, and 'worldly' versions of, supervaluationism.

In this paper, I indirectly argue for supervaluationism and, treating Barnes and Williams's and Akiba's views as representative cases of *modal* worldly vagueness, argue against both the intelligibility of those views and their reasons for preferring them. I will show that there is no reason for preferring the modal treatment of ontic vagueness over supervaluationism, as the former treatment yields no additional explanatory payoff at steep theoretical costs: a recourse to fundamentality, a treatment of precisifications[4] that collapses into semantic facts or epistemicism, and an appeal to vacuous *metaphysical facts* for differentiating between coincident objects or precisified worlds (one where p, and one where ~p).

I will begin Section 1 with a terminological note and some needed background on the types of indeterminacy, supervaluationism, and on the ontic view, before turning to a brief history on some of the latter's best-known opponents (Evans, Lewis, and Russell). Section 2 will explain the modal treatment of metaphysical vagueness through the representative cases, explaining both what the views share, and how they diverge. Section 3 will develop my main argument by isolating three core problems with the modal view: its appeal to fundamentality, its treatment of precisificational space, and its dependence on metaphysical facts. In Section 4, I respond to the representative cases and against their arguments for the intelligibility and favoring of their views. Finally, I will consider some objections to both my arguments and my arguments' responses in Section 5, before ending with some brief remarks in the form of a conclusion.

SECTION 1: TERMINOLOGY, BACKGROUND, AND TYPES

Before turning to the various types of indeterminacy, a brief terminological note is in order. Many theorists dealing with vagueness seem to make little distinction between the phenomenon itself and the typical referent of 'indeterminacy'. As counterexample, Eklund fears that the lack of distinction blurs concepts which are otherwise distinct: vagueness is a phenomenon that "centers on the sorites paradox" (Eklund 2011, p. 354) while much of what is referred to as indeterminacy (e.g., the open future, quantum indeterminacy) does not. In what follows, I will draw no such distinction and, as most other theorists, use terms such as 'vagueness' and 'indeterminacy' interchangeably[5].

The philosophical literature covers three main sources of vagueness: semantic, epistemic, and either ontic, objective, metaphysical, or worldly (also referred to here interchangeably). Largely, this paper is concerned exclusively with the semantic and worldly types of vagueness. But before putting the epistemic kind of indeterminacy aside, a brief explanation is needed. Epistemic vagueness (or, in this context, 'epistemicism') is the view that treats vagueness as reducible to matters of knowledge; more specifically, to a lack of knowledge as the source of some or all vagueness. Viewed differently, epistemicism about some vague proposition p entails that the world is in some way *determinate* regarding p, but that the source of p's vagueness is a matter of epistemic status or accessibility. While epistemicism about vagueness is perfectly legitimate[6], it is a minority view that seems partly ignored by philosophers in the vagueness literature. I will continue with this trend, leaving much that could be said about the view aside.

Unlike epistemicism, the semantic or linguistic view is a majority view which suggests that the source of vagueness is either in our language, or more broadly in the representational features of the concepts and expressions we use to describe the world. While it would be difficult to deny the presence of *any* vagueness in our language, explicit proponents of the semantic view often maintain that the source of *all* vagueness is representational and are thus unlikely to endorse other views alongside their own, since those views would be mutually incompatible (Conversely, the relevant proponents of the worldly view, i.e., those discussed in this paper, don't necessarily deny the existence of *some* semantic indeterminacy, but rather think of neither worldly nor semantic indeterminacy as fully exhaustive[7]). Assuming that the notion of semantic vagueness is intuitive enough to not need explanation, I will move onto the specific semantic view which is of most relevance to this paper: supervaluationism[8].

Fine popularized supervaluationism as an alternative framework to the possible treatment of "neither-true-nor-false, or Indefinite, as a third truth-value" (Fine 1975, p. 267). Unlike the non-classical, i.e., three-value approaches that Fine is motivated to avoid, in which logical connectives such as '&' become non-truth-functional[9], the supervaluationistic framework preserves logical relations, so that 'P & R' (for the mutually exclusive predicates 'pink' and 'red') is false, even though P and R are both borderline cases. This is made possible via a specification space of admissible interpretations that Fine calls *precisifications*. These precisifications are all the semantically admissible ways in which a vague sentence can be made precise and thus be true or false. As such, a sentence is indeterminate in truth value iff it is true in *some* admissible interpretations and false in others, where an admissible interpretation assigns to the term a precisification (Alternatively, a sentence which is true under *all* admissible interpretations is determinately true and called a super-truth, hence the term 'supervaluationism'). From all of this, it follows that standard supervaluationism is assumed to reject bivalence but adhere to excluded middle. Lastly, while the semantic view does not entail supervaluationism (or rather, while one need not be a supervaluationist to be a semanticist about vagueness), the supervaluationistic approach provides an effective framework for dealing with cases of indeterminacy while preserving meaning in vague language. As such, supervaluationism is a widely held position among philosophers who adhere to the linguistic/semantic majority view.

Having addressed both epistemic and semantic views, we are left now with the worldly, metaphysical, objective, or ontic view, which suggests that the source of all or some indeterminacy is the world itself. Now, in an effort to explain the view, I can't help but notice the surprising difficulty by which one can say more about it[10], for the view's main challenge is seemingly its lack of (reductive) definition. Without jumping ahead to the point where I give details on the relevant philosophers' attempts at (or rejections of) reductive definition, I will give Barnes', for at the moment such an attempt seems the closest thing available:

- (OV) Sentence S is ontically vague iff: were all representational content precisified, there is an admissible precisification of S such that according to that precisification the sentence would still be non-epistemically indeterminate in a way that is Sorites-susceptible (Barnes, 2010 p. 604).

In other words, this is a way to negatively define ontic vagueness by saying that it is what remains when sentences are completely precisified, and yet continue to be non-epistemically indeterminate. Other defenders of metaphysical vagueness give similar or no sufficient-and-necessary-condition-style definitions altogether. This includes Rosen and Smith (2004), Taylor and Burgess (2015), and Akiba (2000 and 2004). In fairness to these philosophers, rather than making fully affirmative claims about the existence of ontic vagueness, almost all of them (with the exception of Akiba) seem to be responding to the claim that metaphysical vagueness is itself unintelligible or incoherent. Nevertheless, I'd argue that the reason these philosophers omit sufficient-and-necessary-style definitions is a shared assumption that the phenomenon is itself either intuitive or irreducible/fundamental, and thus immune to or evasive of definition. I will say more about this in Section 3, where I highlight core problems with the ontic view and its modal treatment.

Before concluding this section, it is worth noting three notable challenges to the worldly view. Beginning chronologically, there is Russell (1923), who very plainly accuses the worldly view of projecting linguistic and symbolic properties onto the world itself, and of committing the 'fallacy of verbalism' by confusing the properties of words for the properties of things. Also, there is notably Lewis, who in Section 4.3 of On The Plurality of Worlds defends unrestricted composition by suggesting that restriction (i.e., restricted composition) cannot be accepted, for it is obvious that there is vagueness, and if there is vagueness *and* restriction then there is restricted composition, and therefore it is a vague matter whether composition takes place. Lewis also asserts that "the only intelligible account of vagueness locates it in our thought and language", and therefore "vagueness is semantic indecision" (Lewis 1986, p. 212). Finally, there is the (perhaps) most referred-to challenge to metaphysical vagueness, Evans's (1978) who with a single-page paper, presents a formalized reductio against the existence of vague objects. Evans's argument goes: 'a' and 'b' are singular terms such that the sentence 'a = b' is of indeterminate truth value (let's call this, the 'indeterminate-identity-statement' assumption); this means that, in reference to b, we can now introduce the property of it 'being indeterminate whether it is a'. But then we have a = a, which shows that a *cannot itself* possess the property of it 'being indeterminate whether it is a'. And so, from Leibniz's law we derive that ~(a = b), for b possesses the property of 'being indeterminate whether it is a' and a does not (and thus there is a property they do not share); therefore, there is a contradiction in the starting assumption that the sentence 'a = b' is of indeterminate truth value (i.e., in our 'indeterminate-identity-statement'). While Russell's and Lewis's challenges aren't entirely relevant to what remains of this paper, a lot can be said about Evans's argument and its impact. I will stop here, however, before briefly revisiting Evans in my discussion of Akiba.

SECTION 2: THE 'MODAL' TREATMENT OF ONTIC VAGUENESS

This section treats Barnes and Williams and Akiba as representative cases of a modal approach to metaphysical vagueness. Their views differ in important respects, but both understand ontic vagueness by appealing to precisificational space and special modalities. But before describing the overlapping and unique features of the relevant modal treatments, it is appropriate to clarify the respective philosophical projects of the chosen representative cases. Firstly, Barnes's (2010) project seems explicitly intended as a response to accusations of unintelligibility and incoherence on behalf of those unfriendly to metaphysical vagueness. Barnes and Williams (2011) is no different, but seems marginally less direct in its response to such claims; instead, it provides a general theory of metaphysical indeterminacy (or 'indefiniteness', as it is referred to). The point of providing such context is the following: neither Barnes (2010) nor Barnes and Williams (2011) strongly or straightforwardly assert the existence of metaphysical vagueness; rather, these philosophers provide a working theory for the ontic version of indeterminacy. Akiba, on the other hand, seems explicitly motivated to "defend a version of the worldly view of vagueness" (Akiba 2004, p. 408), arguing for the preference of a modal treatment over the semantic (i.e., supervaluationist) alternatives. Unlike Barnes and Williams, I take Akiba to be making a much stronger claim about what *is*, and a normative claim on which view those sympathetic to supervaluationism ought to prefer. The point is that these are relevant distinctions that frame both the extent to which the aforementioned philosophers respectively succeed and the forthcoming arguments.

Having clarified the projects of the representative cases, I will now say a few words about what each view has in common: what makes the representative cases *modal views* of metaphysical vagueness, and why I am lumping them together. In brief, both Barnes and Williams' and Akiba's views share two distinct features: 1) a primitivist handling of vagueness, and 2) a modal treatment of supervaluationism and indeterminacy, where the corresponding space of precisifications is treated as a space of possible worlds. Past these broad descriptions, there are significant differences between both views (see Akiba (2015) for contrast and a criticism of Barnes and Williams (2011)). But before moving onto these differences, (2) deserves emphasis, for it makes the understanding of their views as modal views much clearer.

In short, both views are modal views insofar as they treat metaphysical vagueness as a sort of determinacy modality similar to alethic modality (possibility, necessity), where, in the first case, 'determinately' is formally analogous to 'necessarily' on the latter. With this treatment comes a corresponding space of possible worlds, where in the case of determinacy the space is that of precisifications (hence the modal treatment being supervaluationistic), rather than that of alethic possible worlds. From this point on, there are material differences between how each view deals with their respective world-space, and between their treatments of precisifications as world-like. But before saying something about this, I'd like to say something about how these two notions of precisification overlap: in that, according to the general modal approach, the world is vague, and the vagueness of the world is some sort of *unsettledness* across competing precisifications.

Finally, it is my belief that these modal approaches are far from ontologically straightforward, as they postulate a world-space structure that raises questions about what these objects *actually* are. Nevertheless, I will move onto explaining each view individually, hoping that doing so will bring clarity to the ontology.

SECTION 2.1: BARNES AND WILLIAMS

To describe the Barnes and Williams' version of ontic vagueness ('BW' from now on), I will highlight what I believe to be its most relevant and distinct features: 1) the assumption that, unlike with epistemic or semantic vagueness, it isn't possible to provide a reductive analysis or definition of metaphysical indeterminacy in fundamental or otherwise more basic and accepted terms, for BW's primitive notion of metaphysical indeterminacy is such that vagueness *is* fundamental, 2) the preservation of classical logic and the aim for fully bivalent truth, where for any p either p or ~p is true and what is indeterminate is which model (i.e., which world) is the intended one, and 3) the strong analogy between BW and both alethic modality and the supervaluationist view, whereby BW postulates some type of abstract, non-alethic possible worlds to serve the role of precisifications.

Regarding (1), and on the need for a biconditional (i.e., necessary-and-sufficient-condition style) definition of metaphysical indeterminacy similar to that which *can* be given for the epistemic and semantic views of vagueness, BW suggests that "it is natural to expect the friend of metaphysical indeterminacy to provide an analogous biconditional, setting out what metaphysical indeterminacy reduces to"; but that "there is an alternative. She may simply deny that facts about metaphysical indeterminacy are reducible to any more basic facts—i.e., that facts about metaphysical indeterminacy (if there are any) are fundamental" (Barnes and Williams 2011, pp. 3-4). Their position, they claim, is equivalent to that found in the metaphysics of time and modality, where contingency and tense are taken as fundamental and irreducible. From this, an important admission of BW follows: that in order to "parallel the [epistemic and semantic] biconditionals, [BW] would have to identify a reductive basis of recognizably metaphysical facts, and then try to characterize necessary and sufficient conditions for m-indefiniteness in terms of what such facts obtain". But while one can "have a rough sense of what semantic facts are (facts about language-use, meaning, and so on) and what epistemic facts are (facts about knowledge, ignorance, justification, and so on), it is hard to know what to make of a particular kind of distinctively metaphysical facts" (Barnes and Williams 2011, p. 4). More on this will be said later.

Regarding (2), BW (2011) notes that, unlike other accounts of indeterminacy governed by nonclassical logic where excluded middle fails, "the account to be developed [there]… [is] fully classical and bivalent" (Barnes and Williams 2011, p. 8). While this introduction serves as a setup for outlining how such classical and bivalent logic is secured, the rest of Section 2 provides a necessary framework for their (i.e., BW's) modal treatment of metaphysical indeterminacy, and thus for (3), which is what the remainder of this subsection will be about.

BW assumes that metaphysical indeterminacy consists in a fundamental kind of unsettledness in the world; and thus, when p is metaphysically indeterminate, there are two possible (exhaustive, exclusive) states of affairs: the state of affairs that p and the state of affairs that not-p, and it is simply unsettled which in fact obtains. This primitive conception is such that it allows for no third state of affairs: the state of p being indeterminate. How BW chooses to transition from here and into a modal conception is as follows: by postulating a world-space consisting of ersatz world-like (albeit, abstract) objects which represent ways the world *might* be. Of these representations, i.e., of the postulated ersatz world-like objects, there is potentially one which is actualized, representing things exactly as they are. It follows then, that there is some way to carve up said world-space so to "single out a new modality—the worlds which are precisificationally possible" (Barnes and Williams 2011, p. 9), where a world is such if it does not determinately misrepresent reality. This entails that, if the actual world doesn't contain *any* indeterminacy, then there is one world in BW's world-like precisificational space; but if the actual world isn't determinate, then there is more than one candidate world for the *actual* world, neither which is determinately correct. From here, BW uses this precisificational space similarly to that of other precisificational indeterminacy theories:

- for any p, p will be determinate if all the worlds in the space of precisifications represent that p
- p is indeterminate if some worlds in the space of precisifications represent that p and some represent that ~p

Finally, and for the purposes of constructing the relevant world-space, BW suggests that "when matters are metaphysically indeterminate, it is indeterminate which world is actualized—and hence it will not be settled which of p or ~p is True, though one or the other must be. In a certain sense, then, we can agree that there is a precise way that things are—so long as by 'precise' one means that for every p, either p or ~p, and either p is True or ~p is" (Barnes and Williams 2011, p. 10).

SECTION 2.2: AKIBA

Similarly to BW, I will now highlight what I believe are Akiba's (2000 and 2004) most distinct features: a) while (2000) provides a definition-*like* description of vague objects, Akiba provides neither a clean nor an explicit necessary-and-sufficient-conditions-style definition of worldly indeterminacy (2000 and 2004), b) without explicitly using terms such as 'fundamental' (or 'primitive', etc.), Akiba not only endorses a non-reductive view similar to that of BW, but goes one step further by thinking of vagueness as *pervasive* (my term), where even "microscopic objects such as cells and molecules, are vague to some degree" (Akiba 2000, p. 359), and c) a different treatment of vagueness-as-modality from that of BW, with an even stronger analogy to supervaluationism, where Akiba's possible worlds are precisifications in a four-dimentionalist-type space, while vague objects are transworld and potentially coincident objects.

No further description of (a) seems needed, for this attribute should be self-explanatory. (Separately, I will return to this definition-like description of vague objects later.) On (b), Akiba notes that "one implication of [the] view is that vagueness is not completely explicable in precise terms" and that his intention is not "to give a reductive account of the former in terms of the latter. On the contrary, [the] working hypothesis is that the world is vague, through and through. Even when we consider microscopic objects such as cells, molecules, and atoms, we find vagueness in their properties and boundaries" (Akiba 2004, p. 423). Furthermore, Akiba makes a similar (albeit, quantificationally-unspecific) claim in (2000) when he notes that "familiar examples of vague objects are medium-sized objects with indefinite spatio-temporal boundaries, such as geographical areas, furry animals, and persons, but on this view almost all objects, including microscopic objects such as cells and molecules, are vague to some degree" (Akiba 2000, p. 359).

Regarding the most important of the relevant features, i.e., (c), Akiba describes his worldly (or objective, or ontic) view as one that "exploits a strong analogy between vagueness on the one hand and metaphysical and temporal modalities on the other", and as one that has "a structural similarity to" or "may even be considered to be the worldly version of supervaluationism" (Akiba 2004, p. 408). To develop the temporal modalities comparison further, Akiba explains the modal conception of vagueness in terms of indefinite temporal four-dimensionalism, where "temporal slices of the world may be taken as possible worlds of a sort (i.e., temporal worlds)" (Akiba 2004, p. 408). Similarly, and on Akiba's view, "the entire world has another dimension: the precisificational dimension… which concerns determinacy and indeterminacy, [and] is made up of precisified worlds, in which everything is precise" (Akiba 2004, p. 408). So, to summarize, Akiba's ontology includes world-like objects which play the role of supervaluationist precisifications and constitute a fifth modal dimension similar to that of time and/or possibility/necessity. Now, I will take a moment to clarify: per (2015), Akiba seems to be an *actual* realist about these precisified possible worlds (or world-like precisifications), suggesting that according to his ontic view "a possible world is just like reality (or the actual world), consisting of concrete and often material entities such as individuals, events and states of affairs" (Akiba 2015, p. 567). The extent to which Akiba is committed to this type of realism in both 2000 and 2004 seems less clear. He does, however, leave the question explicitly open in (2004, p. 408), which suggests options in how we ought to interpret the reading. Given that it would be uncharitable to interpret Akiba in the strongest possible sense (i.e., as a realist), I will assume a fully open and uncommitted view for the purposes of a proper dialectic.

Finally, and with regards to the portion of (c) about vague objects and how they populate Akiba's fifth-dimensional precisified worlds, the view suggests that said objects are transworld objects that extend over precisified modal space, similarly to how ordinary objects extend over three-dimensional space. In this view, vague objects may coincide with each other in one world and diverge in another: they are indiscernible in a world when they have all the same properties in said world (i.e., they overlap completely). Now, with regards to this, Akiba makes a key distinction: coincidence (i.e., indiscernibility) in two worlds isn't sufficient for identity, for strict identity is indiscernibility in *all* precisified worlds. According to him, other philosophers have treated the two (coincidence and identity) to be the same, but that is not the view Akiba advocates (At this point it may be worth noting that Akiba's treatment of identity is directly motivated by and developed as a response to Evans, for Akiba accepts the validity of the Evans (1978) argument but denies the assumption that the existence of vague objects entails the truth of indeterminate identity sentences (e.g., 'it is indeterminate that a = b'). Furthermore, Akiba's view suggests that "no claim of strict identity is indefinite in truth-value, but some claims of coincidence are, if they refer to vague objects" (Akiba 2000, p. 361). Evans's conclusion isn't inconsistent with the claim that a and b (*can*) coincide, so strict identity claims need not be indefinite; but there can be indefinitely coinciding objects, and both of those claims aren't inconsistent with each other (or so Akiba maintains). Importantly, it may be worth noting that Akiba's claim is not that there are such vague objects, but rather that despite Evans's argument, vague objects are *possible*).

SECTION 3: PROBLEMS WITH THE MODAL VIEW AND REPRESENTATIVE CASES

The objections that follow are aimed at the modal treatment of metaphysical vagueness as represented by the representative cases. Since both sets of views share a modal approach, the problems I identify here are aimed at their shared strategy. But before outlining what I believe to be the core issues with the relevant modal treatments, I will present a sketch of a rough and basic argument for my main thesis and view:

- P1: For two equally intelligible and competing views V1 and V2, there is no reason for preferring V1 over V2 if V1 is similar enough to V2 in all relevant explanatory ways except if V1 has problems that V2 doesn't have.
- P2: The relevant modal views of ontic vagueness are similar enough to supervaluationism in all relevant explanatory ways but have problems that supervaluationism doesn't have.
- C: Therefore, there is no legitimate and meaningful reason for preferring the relevant modal treatments of ontic vagueness over supervaluationism

Given its obvious simplicity, the argument is valid, so what needs argument is whether the premises are true. P2 is clearly the one that needs evidence or argument, while P1 is less controversial; however, I'd like to say a few words about P1 not to ignore a glaring weakness: what do I mean by 'has problems'? In short, I mean the intuitive, pre-theoretical way in which a view or theory has 'problems': things like explanatory gaps, contradiction, unconvincing assumptions, circularity, or similar. The aim is to keep problems (or weaknesses, or whatever) somewhat loose and open, and to appeal to an uncontroversial notion of parsimony throughout the argument. I will say more on this later but will defend P2 in the way of evidence (and some argument) in what immediately follows.

To begin, I'd like to group the relevant issues with each set of modal ontic views into three shared core problems:

- lack of definition(s) and appeal to brute, fundamental facts
- the *admissible interpretations* gap
- the implicit or explicit appeal to metaphysical facts

Each of these problems can now be taken in turn.

SECTION 3.1: FUNDAMENTALITY

Before attempting to deny the treatment of metaphysical vagueness as fundamental, I need to say a few things about how the representative views do such a thing. Firstly, Barnes (2010) gives a negative definition (OV), while BW (2011) simply asserts metaphysical vagueness as fundamental while appealing to intuitive notions of indefiniteness. Akiba on the other hand claims that "the world is vague, through and through" (Akiba 2004, p. 423). While the former two treatments (Barnes (2010) and BW (2011)) are more clear appeals to fundamentality, the latter is less so, for to say that metaphysical vagueness is fundamental is to say it cannot be explained (in other, more general or simpler terms) while to say that metaphysical vagueness is pervasive (in the world) is to say something about a feature of the world. But here is why I am choosing to couple the notions together: because in the case of metaphysical vagueness, and in the absence of a clear story on instantiation or distribution, you need one notion in order to make sense of the other. I will explain.

First, let metaphysical vagueness be fundamental but *not* pervasive. Then, some things (presumably objects) are vague and others aren't, to which the proponent of metaphysical vagueness owes an instantiation story or explanation. But the proponent of fundamental metaphysical vagueness *won't* provide an instantiation story or reductive explanation, because any valid story or explanation must appeal to other base level or independently motivated facts, which the proponent of metaphysical vagueness wouldn't have: they believe vagueness is irreducible and fundamental. Furthermore, any *possible* story about distribution or instantiation must explain epistemic access, for the proponent of metaphysical vagueness must explain how she knows to discern between cases: vague vs. determinate. So, to summarize, the proponent of non-pervasive metaphysical vagueness must have a story about instantiation and distribution, and an explanation of epistemic access, all of which BW seems to lack when appealing to fundamentality. This won't do.

Now, let metaphysical vagueness be fundamental *and* pervasive; it follows that *all* things, including the most fundamental simples, are metaphysically vague. But this suggests that nothing is determinate, which in turn entails the exclusive disjunction of contingent pervasive indeterminacy or necessary pervasive indeterminacy. On the first disjunct, if contingent pervasive indeterminacy is true, then there should be possible worlds where metaphysical vagueness isn't pervasive, which leads to the question: how do we know our world isn't such a world? The only conceivable (or admittedly, acceptable) answer would be to appeal to empirical (i.e., epistemically accessible) facts, since we must be able to point to some physical empirical fact to explain how we know. But this makes the relevant type of vagueness non-fundamental, for it means that it *can* be explained in other, more generally accepted terms (i.e., physical empirical terms), and that will not do. We are then left with the second disjunct: necessary pervasive indeterminacy, which itself entails another disjunction. Either the necessary indeterminacy is nomologically necessary, metaphysically necessary, or logically necessary. Of the first option, we can say of nomologically necessary indeterminacy the same we said about contingent pervasive indeterminacy: it should be explainable in physical empirical terms. This leaves only metaphysical or logical necessity. Of both, however, it can be said that fully determinate objects are perfectly conceivable; and while conceivability isn't possibility, I don't see why fully determinate objects should be metaphysically or logically impossible[11]. Here, then, is our first core problem: for any combination of fundamentality and pervasiveness, the modal view must deal with *either* the explanatory burden of instantiation, distribution, and the problem of epistemic access, *or* the assumption that everything must to be vague. In turn, the latter assumption either collapses into some physical theory about the world or commits the modal theorist to the claim that everything being vague is logically or metaphysically necessary, which, as I've shown, is easily deniable by appealing to conceivability.

SECTION 3.2: ADMISSIBLE INTERPRETATIONS

As mentioned, Fine (1975) presented supervaluationism as an alternative framework to third-value logics, where rather than treating indeterminate as a third truth value, vague sentences become true or false relative to admissible interpretations in a specification space of precisified candidates. These precise possible interpretations Fine calls precisifications. In the modal versions of metaphysical vagueness, a similar theoretical apparatus is employed, where precisifications either play the role of, or simply *are*, possible worlds. In the semantic view, precisifications are explained in a context of admissible interpretations, which can be understood solely in semantic terms. And here, in the form of a question, is the modal view's second core problem: how, according to its proponents, are we to make sense of their space of admissible interpretations, or, i.e., their space of precisified worlds? Taking both views separately, I don't see how Akiba's version could avoid appealing to semantic facts in making sense of its fifth dimension and world-space; in fact, I'd say an explanation on the composition of its world-space is missing, and so a story of how we are to define it (i.e., said space of worlds) is owed. And for BW I say the same and more. Their view couples the same semanticist problem with the problem of defining indeterminacy in their world-space as *unsettledness* over which a precise candidate world obtains; but how are we to make sense of this unsettledness other than in epistemic terms? Likely, their response is that they do not need to make sense of it, for they are treating unsettledness as metaphysical vagueness, and metaphysical vagueness is irreducible to more generally accepted or simpler terms, to which I respond by pointing to the first problem above. In circles we go.

SECTION 3.3: APPEAL TO METAPHYSICAL FACTS

Let's credit and grant that, in the context of irreducibility and the absence of definition (i.e., the biconditional reduction), BW (2011) acknowledges the issue outright:

> On our favoured view, the putative subject matter of metaphysical indeterminacy is strikingly disanalogous, in this respect, to that of either epistemic or semantic indeterminacy. To parallel the earlier biconditionals, we would have to identify a reductive basis of recognizably metaphysical facts, and then try to characterize necessary and sufficient conditions for m-indefiniteness in terms of what such facts obtain. While we have a rough sense of what semantic facts are (facts about language-use, meaning, and so on) and what epistemic facts are (facts about knowledge, ignorance, justification, and so on), it is hard to know what to make of a particular kind of distinctively metaphysical facts. Perhaps this is one source of the scepticism one finds towards the notion of metaphysical indeterminacy in much of the contemporary literature. (Barnes and Williams 2011, p. 4)

But while the acknowledgment attempts to address the issue in the context of irreducibility, it does not, neither in the passage nor elsewhere, address the question of what kinds of metaphysical facts make two candidate worlds, one where p and one where ~p, different from each other, for those differences could be neither semantic or epistemic on the one hand, nor physical or empirical on the other[12]. Of semantic or epistemic facts, I assume no lengthy explanation is needed, since it should be obvious that explaining the candidate-world differences in semantic or epistemic facts would collapse BW into such facts. Of the physical side, less is clear; but this is what I have in mind as a possibility: if BW postulates abstract possible worlds, then I take candidate worlds to be physically identical, for if they were physically distinct then it would be possible to conceive of one candidate world representing the actual world *closer* than all others, in which case that would be the actual world (i.e., there wouldn't be candidates). In fairness to BW, the view could imply the following: candidate worlds may be physically distinct in equal-candidate kinds of ways, which we should expect because the world is vague and there are no physically determinate ways to represent it. But I *don't* think this is what the BW theorist has in mind. I think the BW theorist believes that the differences must amount *only* to metaphysical facts, for they have not only admitted so[13], but they have postulated candidate worlds as being fully precise, and so admit determinacy is possible. Most importantly, I think that the most common examples cited by proponents of metaphysical vagueness (with the exception of perhaps those who, like Akiba, think that indeterminacy is pervasive) cohere with this notion: the reason it is indeterminate whether a certain atom belongs to Kilimanjaro is because Kilimanjaro is a mountain with imprecise boundaries; but it wouldn't be right to say the location of the atom isn't precise[14], for that would counter common sense notions of physical facts. So, if the location of the atom isn't imprecise, then there is a way to represent its relation to Kilimanjaro (or rather, to the rest of the atoms around the object) precisely, and there should therefore be at least one candidate world representing this relation precisely, and so on. Now, the defender of BW (or Barnes or Williams themselves) could respond by saying that 'it is possible for representations to be perfectly determinate, but not possible for the world to be physically precise', to which I say a story or explanation is owed. While I hope the directly aforementioned suffices for making the third core problem obvious, let me note the extent to which similar appeals are applicable to Akiba before going onto this section's final sub-section.

Unlike the BW framework, where it makes sense to speak of distinct abstract worlds for candidate precisifications, it seems appropriate to simplify and think of Akiba's many-world-structure instead as leveraging coincidence and vague transworld objects. With this in mind, we can now transfer the metaphysical-facts-worry using the following question: what facts differentiate between candidate coinciding precisified objects? Like BW, the differences couldn't amount to semantic, physical, or epistemic facts, for that would collapse indeterminacy into said facts. Without extending further, this is the third core problem of the modal view and its representative cases.

SECTION 3.4: PUTTING THE PROBLEMS TOGETHER

In what's preceded, I've highlighted three core problems with the modal treatment of metaphysical vagueness and their representative cases. I would now like to say what this means for the aim of this paper, i.e., for the paper's main thesis.

As recap, here is P2:

- P2: The relevant modal views of ontic vagueness are similar enough to supervaluationism in all relevant explanatory ways, but have problems that supervaluationism doesn't have.

I take all three core problems to be problems that supervaluationism does not have, and I take supervaluationism to be similar enough to the modal view in all generally relevant and/or explanatorily relevant ways. As such, I take P2 to be true; and with P2 being true (and P1 uncontroversial), I take the argument to now be sound.

Having said the above, I'd like to make one brief clarification. Given the argument, I do not need to show that there are further problems with the modal ontic view, and so I don't need the argument to generalize[15]. All I need is to find *some* problems unique to the modal view that aren't present in supervaluationism, for the argument's premises demand nothing more.

Before some final remarks, I'd like to say a few words in Section 5 about some obvious weaknesses of P2. But first, I'd like to respond to the specific positive arguments made in BW (2011), and most importantly in Akiba (2004).

SECTION 4: RESPONDING TO THE ARGUMENTS FOR THE PREFERENCE OF THE ONTIC VIEW

As mentioned in Section 2, both Barnes (2010) and Barnes and Williams (2011) seem intended as responses (albeit, in differing degrees) to the accusation of unintelligibility and incoherence on behalf of those unfriendly to metaphysical vagueness (or at the very least as positive projects for a coherent theory of metaphysical indeterminacy). As such, they both lack direct and affirmative arguments for the existence of said vagueness; and so, there is no argument in favor of this view that I will respond to. What matters instead is their defense against the charge of incoherence.

Let's grant that Barnes (2010) and Barnes and Williams (2011) are each successful in their respective projects: coherent expositions of metaphysical vagueness. Now, and at least for the sake of argument, let's grant P2 and accept that semantic supervaluationism explains all the same vagueness without the three modal problems as outlined above. Then, I'd ask: what is the use for a *coherent* theory of metaphysical vagueness? I can both accept their view as coherent, while challenging that it is in any way preferable to semantic supervaluationism. And so, I will stop here, while I credit their theory-producing attempt and effort fully.

Unlike Barnes (2010) and Barnes and Williams (2011), Akiba is explicit in its effort to "defend a version of the worldly view of vagueness" (Akiba 2004, p. 408), doing so via argument for the preference of modal vagueness over semantic supervaluationism. Akiba's specific claim is that "the modal view has no apparent disadvantage compared to supervaluationism, and indeed it has much advantage. [Akiba] may even go so far as to claim that anybody who is attracted to supervaluationism should embrace the modal view instead" (Akiba 2004, p. 412). Akiba's argument for this claim follows from the idea that both inflationist and deflationist versions of supervaluationism suffer from weaknesses that the modal view does not. For better clarity in my response, I will reconstruct Akiba's argument in clear premise-conclusion form:

- P1: Either you are an inflationist about semantic relations and reference who believes said relations are substantive physical properties, or a deflationist who holds that semantic predicates such as 'refers to' are only logical or disquotational devices.
- P2: If you are a deflationist, you deny the existence of substantive (i.e., genuine, physical, causal) reference relations, and you can't accept supervaluationism, for supervaluationism takes indeterminacy as referential indeterminacy (i.e., indeterminacy in reference relations).
- P3: If you are an inflationist, you take reference relations to be physical relations, and so, you accept indeterminacy as referential indeterminacy in at least some physical relations and thus admit that a specific non-representational part of the world involves indeterminacy.
- P4: But, for the inflationist, it is "extremely odd" (Akiba 2004, p. 412) to hold that just *those* specific physical relations (i.e., referential, semantic relations) involve indeterminacy and no others do: there is no good reason to think that those physical relations are in any way special, for there must be many other indeterminate causal relations in the world.
- C: Therefore, whether you are a deflationist of an inflationist, if you are attracted to supervaluationism, you should instead embrace the modal view.

Granting the validity of the reconstruction, we can say that P1 is trivial, and that neither P3 nor P4 is obviously true. Most importantly, however, I take P2 to be false, and so the argument isn't sound.

To defend this view, I will begin with some minor pushback on P3 and P4, before outright denying P2.

Admitting a possible straw-man interpretation of P3, Akiba's move suggests that inflationists believe in a type of semantic indeterminacy that is substantive in a way that is worldly, and thus should accept worldly indeterminacy, for their view entails it. (See Taylor and Burgess (2015) for a similar everything-is-a-part-of-the-world argument.) While it is possible that this is true of *some* inflationist views, there is much to be said about the indeterminacy that the inflationist is willing to accept; and most importantly about whether said indeterminacy is the same as the one Akiba defends. This point ties into the pushback of P4: granting that the substantive referential relations that the inflationist accepts may be 'worldly' in the relevant sense, why is it 'odd' to think only *that* type of worldly indeterminacy is the worldly indeterminacy that exists? In fact, I think there is good reason for the inflationist friendly to this view of worldly indeterminacy to believe that the only kind of worldly indeterminacy that exists *is* semantic, for they are a supervaluationist who believes we can explain all indeterminacy in semantic terms to start, even if they grant that some of that semantic indeterminacy is worldly.

On P2, there isn't so much to say other than pointing to the fact that a version of supervaluationism where reference relations *aren't* substantive in the relevant sense (i.e., genuine, physical, causal) is conceivable, and thus there are ways to be a supervaluationist and need not accept the argument. Even if it is true as empirical fact that all philosophers who are supervaluationists also believe in substantive reference relations, I can still postulate *n-supervaluationism* (for non-substantive supervaluationism) and deny that P2 is true, since the premise entails one couldn't be a supervaluationist without believing in substantive reference relations, which is obviously false.

SECTION 5: OBJECTIONS

While it may be appropriate to object to my counterarguments (or rather, my responses to arguments) as found in Section 4, Section 3 contains the most substantive of my claims and will therefore be what is discussed in this section. I will, for reasons of brevity, focus mostly on what I expect to be the main objections from said section.

As formulated above, my rough sketch of an argument included these two premises:

- P1: For two equally intelligible and competing views V1 and V2, there is no reason for preferring V1 over V2 if V1 is similar enough to V2 in all relevant explanatory ways, except if it has problems that V2 doesn't have.
- P2: The relevant modal views of ontic vagueness are similar enough to supervaluationism in all relevant explanatory ways but have problems that supervaluationism doesn't have.

While some may object to P1 on grounds of simplicity or ambiguity, I think that a charitable reading should dispel most well-intended worries. For this reason, I'll assume the more obvious objection to be the weakness of P2 is the treatment of supervaluationism as similar enough to ontic vagueness in all relevant explanatory ways. The specific possible objections I will consider are:

1. I haven't shown that supervaluationism is similar enough to ontic vagueness, and/or that
2. I've assumed there aren't problems unique to supervaluationism which also aren't part of the modal view, and/or that
3. I am begging the question by assuming the scope of 'all relevant' to be semantic indeterminacy, i.e., that all in need of explaining is representational, linguistic, etc.

To the first and second objections (which I will deal with in conjunction), I grant them as true; but I will add that I need not accept the associated explanatory burden, for the philosophers I have discussed have *themselves* drawn the comparison and accepted supervaluationism as similar to ontic vagueness in the appropriate ways, and they have not, as far as I can tell, argued against the supervaluationist view by citing some set of possible problems which are unique and exclusive to it. In fact, and as a possible exception to what has been said, Akiba (2004) assumes that supervaluationism must be inflationist, which in fact makes supervaluationism even *more* similar to ontic vagueness than I'd be willing to grant. And, therefore, I don't think the spirit of either objection holds true.

Now regarding the third objection, let me clarify the relevant stance: that I assume that all which needs explaining is semantic indeterminacy, and in this sense there is obviously nothing within the scope of what needs explaining that supervaluationism couldn't explain, and thus there is no ontic gap for the views to be dissimilar in any way, and for ontic vagueness to be different because it explains something that supervaluationism does not. While I wish not to stalemate, I can't help but to respond and suggest a counter response: that the objection is relevant only when assuming an ontic gap which the supervaluationist position would then have to explain. As far as I've shown, the burden appears to be on the proponent of ontic vagueness to show this is the case, and nothing up to this point has convinced me that said proponents would be able to do so.

CONCLUSION

To conclude, I have indirectly argued for supervaluationism, and against the representative cases' arguments for the intelligibility and preference of the modal ontic view. I have shown that there is no reason for preferring the modal treatment of ontic vagueness over supervaluationism, for the views shown via representative cases are similar enough to supervaluationism in the relevant explanatory ways, but have problems that supervaluationism doesn't have (mainly, the three core problems outlined and selected). Of the first problem, I've shown that for any combination of fundamentality and pervasiveness, the modal view must either deal with the explanatory burden of instantiation, distribution, and the problem of epistemic access, or the assumption that everything must be vague, which in turn will either lead the view to collapse into some physical theory about the world, or will trap the modal theorist into suggesting that everything being vague is logically or metaphysically necessary. Respectively, of the second and third problems, I have shown that the parallelling of modal ontic worlds against supervaluationist admissible interpretations highlights an aspect of the modal view which collapses into semanticism, and that both treatments of modal ontic vagueness suffer from similar appeals to metaphysical facts. Finally, I argued against Akiba's claim that the modal view ought to be preferable to those attracted to supervaluationism instead of supervaluationism itself, showing that its supporting premise, i.e., the claim that one need be an inflationist to endorse supervaluationism, is false.

Concurrently, I take the debate dealt with here to be one not fully settleable by argument, for it seems common in metaphysics that theories of the world be colored by our preexisting philosophical and/or pre-theoretical intuitions. While I hope to have shown, albeit indirectly, that there is nothing to be lost by accepting the view that vagueness is representational, questions of worldly determinacy seem especially resistant to anything other than fully intuitive treatment.

End Notes

[1] For mereology and composition, see Merricks (2005) and van Inwagen (1994)

[2] To be addressed later, this paper uses the terms 'vagueness' and 'indeterminacy' interchangeably.

[3] This paper will use the terms 'metaphysical', 'worldly', 'objective', and 'ontic' interchangeably when referring to vagueness, i.e., metaphysical, worldly, objective, or ontic vagueness.

[4] To be explained and defined later in the paper.

[5] For better or worse, this seems appropriate given that the theorists in question do accordingly.

[6] By this I don't mean true, or that epistemicism is a/the view that I endorse. Rather, I mean to say that it is a legitimate view one could have.

[7] This isn't to suggest that the majority of the proponents of the worldly view don't think of it as exhaustive. This is just to say that it happens that the philosophers discussed in the paper happen to not believe this.

[8] Or semantic supervaluationism, or standard supervaluationism; all of which are terms I will also use interchangeably.

[9] This is because with the three-value approach if a colored blob is somewhere on the borderline of pink and red such that it is indefinite that P and indefinite that R, then it is indefinite that P & R; but P & R are contrary predicates (i.e., something can't be both P and R); and, most notably, P & P would be indefinite as well, so & isn't truth-functional with respect to the three truth-values.

[10] Something other than the source of vagueness is the "world itself".

[11] The immediately preceding argumentation style is heavily inspired by Balaguer (2021).

[12] It is perhaps worth noting, that I take equal-candidate BW worlds to be physically identical, for BW takes candidate worlds to be *abstract* possible worlds, and for any physical differences between possible worlds I can conceive of a world that is identical to the actual world, and so a better candidate. This is to say that possible worlds which are physically *different* than the actual world couldn't be (good, fitting, applicable, or similar) candidates, for I can easily think of a better possible candidate: another abstract world that is physically identical to the actual.

[13] See quoted passage immediately above.

[14] Fixing time, that is.

[15] By this I mean, to make some inductive point about other potential problems. I can simply state what I have, and that should suffice.

Bibliography

1. Akiba, Ken. (2000) "Vagueness as a modality," *Philosophical Quarterly* 50, pp. 359–370.

2 Akiba, Ken. (2004) "Vagueness in the world," *Noûs* 38, pp. 407–429.

3. Akiba, Ken. (2015) "How Barnes and Williams have failed to present an intelligible ontic theory of vagueness," *Analysis* 75, pp. 565–573.

4. Balaguer, Mark. (2021) *Metaphysics, Sophistry, and Illusion: Toward a Widespread Non-Factualism* (Oxford: Oxford University Press)

5. Barnes, Elizabeth. (2010) "Ontic Vagueness: A Guide for the Perplexed," *Noûs* 44, pp. 601–627.

6. Barnes, Elizabeth, and Williams, J. Robert G. (2011) "A Theory of Metaphysical Indeterminacy," in: Karen Bennett & Dean W. Zimmerman (Eds), Oxford Studies in Metaphysics vol. 6, pp. 103–148 (Oxford: Oxford University Press)

7. Eklund, Matti. (2011) "Recent Work on Vagueness," *Analysis* 71, pp. 352–363.

8. Evans, Gareth. (1978) "Can there be vague objects?" *Analysis* 38, pp. 208.

9. Fine, Kit. (1975) "Vagueness, truth and logic," *Synthese* 30, pp. 265–300.

10. Lewis, David. (2001) *On the Plurality of Worlds* (Malden, MA: Wiley-Blackwell)

11. Lewis, David. (2004) "Vague Identity: Evans misunderstood," in: Tim Crane & Katalin Farkas (Eds), Metaphysics: A Guide and Anthology (New York: Oxford University Press)

12. Merricks, Trenton. (2005) "Composition and vagueness," *Mind* 114, pp. 615–637.

13. Russell, Bertrand. (1923) "Vagueness," *Australasian Journal of Philosophy* 1, pp. 84–92.

14. Taylor, David E., and Burgess, Alexis. (2015) "What in the World Is Semantic Indeterminacy?" *Analytic Philosophy* 56, pp. 298–317.

15. van Inwagen, Peter. (1994) “Composition as Identity,” *Philosophical Perspectives* 8, pp. 207–220.

The Limits Of Ontological Parsimony: Why Counting Entities Doesn't Settle Metaphysical Debates

Ekaterina Vorozhtcova

Ontological parsimony is a theory-choice principle that sometimes tempts people into ontological conclusions. If two candidate theories fit the evidence comparably well, then, other things equal, the one with fewer ontological commitments is often treated as the better theory. The temptation in metaphysics is to move from that methodological preference to an ontological conclusion: if the leaner theory does not require entities of kind X then we should not endorse X's existence. My aim is not to deny that parsimony can be epistemically relevant. It is to determine the conditions under which parsimony can legitimately function as a reason for preferring one theory over another in a way that bears on what it is rational to accept, and to show why those conditions are often difficult to satisfy in metaphysics. Parsimony is sometimes treated as truth-tracking ("the simpler theory is more likely true"), whereas it is also sometimes treated as pragmatic or aesthetic (simplicity as a virtue for manageability or elegance). Since these roles come apart, a parsimony argument can look stronger than it is unless we specify which role is in play.

I develop this point through the debate about numbers. The debate is ideal for testing parsimony arguments because it invites a tempting inference from dispensability to nonexistence. This inference is not delivered by parsimony alone. At most, it is taken to support a comparative epistemic claim: the more parsimonious theory is more likely to be true, or at least more reasonable to accept. The paper proceeds as follows: In Section 0, I fix working notions of ontological commitment, ontological parsimony, facts/phenomena, and explanatory adequacy, and distinguish parsimony in kinds from parsimony in number (including the question of how kinds are individuated). In Section 1, I state the strongest defensible parsimony principle and separate evidential parsimony from pragmatic simplicity. In Section 2, I isolate the common slide from "not needed" to "does not exist" and show the missing premise. In Section 3, I defend a fixed facts/phenomena constraint on legitimate parsimony comparisons and explain why violating it makes the "tie" claim ill-defined or question-begging. In Section 4, I apply the constraint to the debate about numbers and show where the alleged tie is secured by revising what counts as needing explanation. In Section 5, I conclude that parsimony is a limited-use tool (tiebreaker and burden pressure), not an automatic eliminator, and answer an objection from ordinary reasoning and explain why its justification does not automatically transfer to metaphysical ontology.

SECTION 0: TERMS

This paper is about a very specific argumentative pattern: philosophers often use ontological parsimony (i.e., "fewer entities") to move from a claim about what we need in a theory to a claim about what exists. My aim at this point is to separate several different terms that run together, and then to state the conditions under which parsimony can do genuine epistemic work.

A preliminary clarification: parsimony is sometimes treated as a guide to truth, but it is also commonly treated more modestly as a guide to rational acceptance (or rational suspension) when the available considerations do not decisively discriminate. The stronger truth-tracking reading requires more argument than the weaker rational-preference reading. Many metaphysical parsimony arguments appear compelling because they go from a modest methodological point to a metaphysical conclusion without stating the missing steps.

0.1: ONTOLOGICAL COMMITMENT

A theory is ontologically committed to whatever entities must exist for the theory to be true (This is a working notion, not a deep theory of commitment. The paper does not require settling on the best formal test for commitment).

0.2: ONTOLOGICAL PARSIMONY AS A PROPTERTY OF THEORIES

As Alan Baker puts it: "Ontological simplicity, or parsimony, measures the number of kinds of entities postulated by the theory." (Baker 2022). Ontological parsimony is best treated as a property of theories, not as an aesthetic preference. One theory can be more ontologically parsimonious than another simply because it posits fewer entities. So: If "T1 is more parsimonious than T2" is a descriptive comparison, then "We should prefer the more parsimonious theory" is a further methodological claim that needs defense. Here, parsimony can function as an epistemic consideration in theory adjudication, but it can also function as a merely pragmatic or aesthetic preference for elegance. Since this paper is concerned with ontological conclusions, its focus is on the epistemic role: when, if ever, parsimony provides a reason to prefer one theory over another in a way that bears on what it is rational to accept.

0.3: Parsimony In Kinds vs. Parsimony In Number

There are at least two different measures of ontological parsimony, and they come apart:

a) Parsimony in kinds/types. A theory is more parsimonious in kinds if it posits fewer types of things.
   Example: a theory that posits only physical objects is (in this sense) more parsimonious than a theory that posits both physical and abstract objects (i.e., entities that are non-spatiotemporal and causally inert, such as numbers and sets).
b) Parsimony in number (individuals). A theory is more parsimonious in number if it posits fewer individual entities, even if it posits the same kinds.
   Example: a theory that posits 10 fundamental particles is (in this sense) more parsimonious than one that posits 11, even if both posit only particles.

Most metaphysical debates about numbers are debates about kinds-parsimony (do we need a kind of entity: numbers?), so that is the main focus here. However, kinds-parsimony is sensitive to the individuation of kinds. A theory may look more or less parsimonious depending on whether we count abstract objects as one kind or instead count numbers and sets as different kinds despite both being abstract. So, an appeal to kinds-parsimony is not yet a neutral premise unless the relevant typing scheme is stated and defended. Because the two measures can come apart, it is important to guard against sliding between them within a single argument (for example, treating a reduction in kinds as if it were supported by considerations that in fact bear only on a reduction in number, or vice versa).

0.4: List Of Facts/Phenomena

Ontological parsimony becomes epistemically relevant only in the context of theory comparison. To make any tie claim determinate, we need a fixed list of what the competing theories are being asked to account for. By "facts/phenomena" I mean the claims, data, or constraints that set the target for theory assessment. The idea is that you can't assess whether two theories do equally well unless you specify what they're being assessed on. This is not a metaphysical claim about what facts are ultimate. It is a methodological point about fair comparison.

0.5: Explanatory Adequacy

A theory is explanatorily adequate (relative to a set of facts/phenomena) just in case it accounts for those facts/phenomena in a way that does not leave obvious holes. I will not build a full theory of explanation here. For present purposes, what matters is that we distinguish between what needs explaining and what counts as an adequate response. A theory is inadequate when, instead of explaining a phenomenon, it effectively removes it from the explananda set.

0.6: The Parsimony Principle

Here is a modest parsimony principle (often grouped under "Occam's Razor"): if two theories are assessed against the same fixed facts/phenomena, and if they are genuinely tied (or near-tied) in explanatory adequacy relative to those facts/phenomena, then, other things equal, ontological parsimony can be a reason to prefer the theory with fewer ontological commitments.

The paper's central claim is that in many metaphysical disputes, the tie conditions are rarely met in a way that entitles parsimony to function as evidence. In ordinary contexts, the "other things equal" condition can be met often enough to make parsimony useful.

Section 1: The Defensible Role of Ontological Parsimony

Ontological parsimony has strong intuitive appeal. When two theories are otherwise on a par, it is tempting to prefer the one that posits fewer kinds of entities. At most, parsimony is usually taken to support a comparative epistemic claim: ceteris paribus, the more parsimonious theory is more likely to be true, or at least more reasonable to accept.

The central question: is parsimony genuinely epistemically relevant to theory adjudication in a way that bears on truth or rational acceptance? Or is its role primarily pragmatic or aesthetic, reflecting a preference for simplicity that is not truth-tracking? Those are the live alternatives. My aim is not to deny that parsimony can play an epistemic role, but to specify the conditions under which it can do so and to argue that, in metaphysics, these conditions are often not satisfied in the way parsimony arguments require.

1.1: The Equality Conditions

A necessary condition for epistemic parsimony is that the comparison be well-posed. Ontological parsimony has evidential force for preferring theory T1 over theory T2 only if:

1. Same explanatory task: T1 and T2 are trying to account for the same facts/phenomena.
2. Genuine parity: by shared standards of adequacy, T1 and T2 are tied (or close enough to tied) in how well they account for those facts/phenomena.
3. No hidden purchase: T1 doesn't avoid entities only by taking on other costs. The same standards count against T1.

In many metaphysical disputes, the conditions required to make a genuine tie-claim are rarely satisfied, or else they are satisfied only after the adequacy standards have been narrowed in ways that are not dialectically neutral.

1.2: Structural Point About Interpretation And Ontology

Mark Balaguer's discussion of anti-platonism supplies a structural point that helps locate where ontological conclusions enter. He argues that an interpretive thesis about what a discourse purports to be doing does not, by itself, settle what exists: "the theistic interpretation of ordinary uses of 'God' does not entail theism (i.e., it does not entail that God actually exists)" (Balaguer 2024, p. 176). The missing step is often a truth premise. "S is about X" does not entail X exists; many philosophers hold that if we add that S is true (under standard truth-conditions), then the existence of X follows. The point is that parsimony cannot replace these truth-and-ontology premises; it can at most operate once the comparison frame and the relevant truth-conditions are in place.

Section 1, therefore, fixes two constraints that govern the rest of the paper. First, parsimony is evidential only when the comparison is well-posed: the same facts/phenomena and the same standards of adequacy are held fixed, and adequacy is tied. Second, parsimony by itself supports non-commitment more readily than denial; denial requires an explicit premise of the "evidence-if-exists" type. The next section classifies the main premises on offer in the numbers debate and tests which (if any) can be defended without covert shifting of facts/phenomena.

SECTION 2: THE MISSING PREMISE

This section isolates the logical gap and shows what would be required to close it. The conclusion is not that parsimony is irrelevant, nor that Platonism is true. Any argument that goes from "indispensable in theory" to "nonexistent in reality" relies on a further principle, and that principle carries the real philosophical burden.

I start by writing the argument in its standard step-by-step form. Let K be numbers, i.e., whatever arithmetic discourse purports to quantify over. Platonists take K to be abstract objects. Anti-Platonists deny that there are such abstract objects and replace the Platonist story with alternatives: on nominalist views, the apparent quantification over numbers is eliminable by paraphrase; on aesthetic views, the content of arithmetic is grounded in facts about ideas or cognitive capacities. Let T1 and T2 be competing theories, where T2 posits K and T1 does not. A parsimony argument would appear as follows: T1 accounts for the relevant facts without positing K; T2 posits K; therefore, other things equal, T1 is preferable; therefore, K does not exist. The last step does not follow from the earlier ones. The premises concern what a theory requires to discharge an explanatory task; the conclusion concerns what exists.

The result is that parsimony arguments often look modest while relying on a controversial bridge principle. Once the bridge is made explicit, the dialectic becomes clearer: the central dispute is not about whether parsimony is a theoretical virtue, but about whether existence is tied to evidential expectation, to indispensability, or to some other substantive condition. Where no defensible bridge is available, parsimony most naturally supports a weaker posture than denial. Elliot Sober's formulation makes that restraint precise: one may "slice away" a conjunct "not by denying that H is true, but simply by declining to assert or deny it," and "silence reduces your risk of error" (Sober 2015, p. 71). That is, parsimony supports non-commitment more directly than eliminative metaphysics unless additional evidential premises are secured.

For abstract objects like numbers, the evidential-expectation route is not straightforward. If evidence is construed observationally or causally, the condition fails immediately: abstract objects are not the sort of thing that show up in perception or causal detection. If evidence is construed theoretically, then the issue becomes: would numbers have to play an indispensable explanatory role with respect to a fixed explananda set if they existed? That is precisely what is disputed in the indispensability literature. So, the epistemic bridge premise can be stated, but it cannot simply be assumed.

A second potential premise is metaphysical rather than epistemic. It identifies existence with explanatory indispensability: only entities that are indispensable to a satisfactory account of the explanation exist. On this view, if K is dispensable, then K does not exist. This premise makes elimination easy, but at a cost: it is no longer an argument from parsimony. It builds a substantive existence condition into 'exists': to exist is to be explanatorily indispensable. I do not need to decide whether that criterion is true; I only need to note that adopting it is a substantive step that cannot be hidden inside of "simpler theories." Moreover, the criterion has wide consequences beyond numbers. Anyone who accepts it owes an account of why those consequences are acceptable. Without that, the premise remains an unsupported axiom.

Another diversion appeals to discourse and interpretation. A common starting point is that ordinary mathematical sentences look existential: "There is a prime between 10 and 20," "There are infinitely many primes," and so on. Taken at face value, this language appears to quantify over numbers, and a Platonist may argue that the best explanation of the truth of such sentences is that numbers exist. But that move requires additional premises. In particular, one needs (i) a truth premise (that the relevant mathematical sentences are true), and (ii) a semantic or metaphysical principle connecting the truth of those sentences to the existence of the entities they appear to quantify over (for example, that truth requires appropriate truth-makers). Without those assumptions, an interpretive claim such as "mathematical discourse purports to refer to numbers" does not by itself entail that numbers exist.

Conversely, anti-Platonists often argue that we can explain the usefulness and apparent objectivity of mathematics without ontological commitment to numbers, for example by paraphrasing the apparent quantification away, treating mathematical talk as a useful fiction, or grounding it in structural or inferential roles. But again, this does not yield a negative existential conclusion by itself. Showing that we can do science or mathematics without explicit reference to numbers is at most a claim about dispensability for some theoretical purpose. So, in both directions, the heavy lifting is done not by the interpretive observation alone, but by the substantive principles added to connect discourse and theory-use to ontology.

These options show the methodological outcome. If parsimony is invoked to justify a negative existential claim, then some further premise of the kind just described must be in play, implicitly or explicitly. When the missing premise remains implicit, the argument can look modest and purely methodological while in fact relying on a controversial epistemic or metaphysical bridge principle. Once that premise is made explicit, it becomes clear that the central disagreement is no longer about whether parsimony is a theoretical virtue, but about whether the bridge principle is correct. The disagreement is about whether existence is tied to evidential expectation, or to indispensability, or to a semantic reinterpretation. This also clarifies a common reaction from philosophy of science. It is true that if T1 and T2 are incompatible, they cannot both be true. It is also true that, under an accepted decision procedure, it can be rational to accept T1 and reject T2. In many scientific contexts, parsimony is treated as a legitimate consideration: once rival theories satisfy the same evidential constraints to a comparable degree, parsimony can permissibly influence which theory one adopts for further work. The point I need here is just that preferring one theory on methodological grounds does not by itself show that the entities posited by the rejected theory do not exist.

The conclusion is that parsimony has a legitimate role; it can function as a policy of restraint, and as a tiebreaker once adequacy is secured relative to agreed facts and standards. But conclusions based on conceptual elimination require more. The next sections use the numbers discourse as a case study to show where these additional premises enter and where the facts/phenomena and standards of adequacy are likely to shift, sometimes without being acknowledged.

SECTION 3: WHAT A "WELL-POSED" PARSIMONY COMPARISON REQUIRES

Parsimony can only function as an epistemic consideration if we first have a determinate answer to three questions:

Target: What facts/phenomena are both theories being asked to account for?

Standard: What counts as an adequate account of those facts/phenomena?

Metric: what counts as one "kind" in the kinds-parsimony count?

If any of those are allowed to drift, "T1 and T2 do equally well" becomes unstable.

3.1: Quantifiers Don't Settle Ontology By Themselves

Quine explicitly says "To be is to be the value of a bound variable" is not itself an answer to ontology; it functions as a test of conformity to a prior standard. The point here is that the comparison frame is not self-applying. It must be fixed before parsimony can function as evidence.

3.2: Disputes About Abstract Ontology Often Lack a Stable Adjudication Base

Hartry Field motivates why metaphysical disputes are especially prone to frame instability. He has doubts about two assumptions, including "that the question of what exists has a univocal and non-conventional content" and "that the right way to answer this question is to look at the existential quantifications of our most fundamental theories" (Field 2016, p. P-2). If the very standards for answering "what exists" are contested, then parsimony cannot function as decisive evidence unless the comparison frame is made explicit and held fixed. That sets up section 4, when we apply fixed comparison frame requirement to numbers, we can track exactly where target and standard drift create the appearance of a tie.

Section 4: Numbers And The Instability of "Equality" Claims

Section 4 argues that many anti-Platonist appeals to parsimony in the numbers debate presuppose a genuine tie in explanatory adequacy, but that this tie is often manufactured, under-described, or insufficiently defended.

4.1: THE PURPORT OF MATHEMATICAL DISCOURSE

Starting with a phenomenon: “Our mathematical sentences and theories are about (or at least purport to be about) abstract objects” (Balaguer 2024, p. 176). It is important because it fixes a plausible explanation: mathematical discourse presents itself as quantifying over abstracts. The nominalist cannot simply announce “we don’t need numbers” and treat that as equal adequacy unless the nominalist either (i) explains away this purport, or (ii) argues that it does not belong in the target set. If (ii) is chosen, then the appearance of equal adequacy is achieved by fact narrowing: you make the explanatory task easier by removing a candidate phenomenon from the target list. That may be defensible, but it has to be stated openly. Otherwise, parsimony is being used as if it were deciding between theories that face the same fixed explananda, when in fact the explananda have been revised to secure parity.

4.2: THE INDISPENSABILITY DIALECTIC

Indispensability arguments are often presented as if “mathematics is indispensable to science” were a single, unambiguous claim. But there is an important ambiguity in what “indispensable” is supposed to mean. Sometimes it means indispensable for formulating or regimenting a scientific theory (writing it down, deriving predictions, applying it in modeling). Sometimes it means indispensable to the explanatory content of the theory itself. Daly and Langford (2009) highlight this distinction with a caution that is easy to miss in faster presentations of the indispensability argument: “Platonists must show that mathematics has a genuinely explanatory role to play in science. It is not enough to show that mathematics is indispensable to the formulation of scientific theories” (Daly and Langford 2009, p. 644). This point can be made without committing to any verdict about particular scientific examples. The thought is simply that indispensability for using a theory does not automatically settle what is doing the explanatory work within the theory. A theory might require mathematics as representational scaffolding even if, at the level of explanation, what is doing the work is entirely physical. A simple non-technical illustration is the difference between a label and a cause: if I say, “the 17th trial was the first to exhibit outcome A,” I have used a number indispensably to locate an event in a sequence, but the number 17 is not part of what explains why A occurred. It is a way of picking out which event we are talking about. Daly and Langford’s distinction is that some uses of mathematics in science might function in this "tracking/labeling/indexing" way rather than as part of the explanatory story. If one side treats formulation-indispensability as sufficient for ontological commitment, while the other treats only explanatory-indispensability as sufficient, then equal adequacy is not a stable claim unless the relevant adequacy standard is fixed. In other words, the disagreement is not merely over whether mathematics appears in our best theories, but over what role that appearance is being asked to play in the inference from science to ontology.

4.3: EXPLANATORY STRATEGY

Daly and Langford describe a strategy associated with Mark Colyvan (and Alan Baker) as an attempt to meet the stronger demand, namely to argue that mathematics is not merely present in scientific theories but is genuinely part of scientific explanation: "Mark Colyvan and Alan Baker have sought to defend Platonism… that mathematics does play an indispensable explanatory role in science" (Daly and Langford 2009, p. 642). The idea is that if mathematics is doing explanatory work, then it is harder to treat mathematical entities as dispensable scaffolding.

The central pushback is the indexing reply. Daly and Langford formulate it like this: "Even granted that the formulation of these explanations involves indispensable mention of mathematical entities, it does not follow that these entities play a part in the explanations" (Daly and Langford 2009, p. 645). The point is not that mathematics is absent from the statement of the explanation. It is that presence in the statement does not automatically amount to an explanatory contribution. On this view, mathematics may be indispensable to how we represent the explanation without being an explanatory ingredient.

Daly and Langford then apply this to Colyvan's specific case and claim that the step to explanatory indispensability is not supported there: "Colyvan provides no reason to think that mathematics has more than an indexing role to play in this example" (Daly and Langford 2009, p. 646). You do not need to accept their example-by-example verdict in order to use this dialectic for your purposes. The methodological payoff here is conditional; if there is a meaningful distinction between formulation indispensability and explanatory indispensability, then the move from "mathematics is indispensable to science" to "therefore mathematical entities exist" depends on which notion of indispensability is doing the work and on what is being counted as an adequate explanation.

That conditional point is enough to support the paper's main point. In the numbers debate, apparent ties are often announced too quickly. Whether a nominalist and a Platonist "do equally well" depends on whether explanatory contribution is part of the target and on what counts as a genuinely explanatory role rather than a merely representational one. If those standards are not fixed explicitly, then the "tie" premise is unstable, and parsimony cannot legitimately function as the evidential discriminator it is often taken to be.

4.4: Why The Dispute Is Not Easily Stabilized

Field shows why the numbers debate doesn't behave like the clean cases where you say, "two theories tie, then parsimony breaks the tie." For one thing, he points out that people aren't always even aiming at the same target when they argue about "Platonism." "'Platonism' can mean a number of things," and he says his earlier use was "a thesis about what exists" (Field 2016, p. 3). That matters because it's easy for one side to be making a metaphysical claim about ontology while the other side is really making a methodological claim about what we should accept given the role mathematics plays in science.

He also pushes back on a very natural thought: that we can settle ontology just by looking at what our best theories quantify over. He doubts "that the right way to answer [the question of what exists] is to look at the existential quantifications of our most fundamental theories" (Field 2016, p. P-2). You don't have to take Field's whole view on board to see the pressure point; if one camp treats quantification as decisive for ontology and the other treats it as part of how we regiment or represent the theory, then the dispute isn't sitting on a shared standard of assessment.

That's why the numbers case is so prone to *frame instability*. In lots of ordinary scientific disputes, the data and success conditions are relatively fixed. Here they aren't. People disagree about what counts as explanation, about whether mathematics is explanatory or merely representational, and about how to read ontological commitment of a theory's formulation. And if those standards aren't settled, then the "they're equally adequate" premise you need for a parsimony argument isn't settled either.

4.5: Strong Tie And Weak Adequacy

Now we can state the main conclusion of Section 4. The utterance “we can do everything without numbers” is ambiguous:

> Strong tie claim: under fixed explanation and fixed standards, the nominalist matches the realist in explanatory adequacy.
> Weak adequacy claim: the nominalist matches only after narrowing the explanation or thinning the standards.

Only the strong tie claim would license parsimony as evidence. If the nominalist only achieves weak adequacy, then parsimony cannot do eliminative work; it can only press the realist for justification of additional ontology and support a policy of restraint.

Conclusion: Parsimony as Burden Norm, and Restraint

Ontological parsimony can have epistemic force only within a fixed comparison frame. When a genuine tie (or near tie) is established under fixed explanation, fixed adequacy standards, and a stable kinds metric, parsimony can rationally favor the leaner theory. But in metaphysical disputes about numbers, the tie premise is often unstable because the explanation and adequacy standards are themselves contested.

Field’s doubts about the Quinean framework help explain why both easy eliminations and easy indispensability inferences can misfire when used without methodological discipline. Daly and Langford, meanwhile, show why the realist burden is not discharged merely by showing that mathematics is useful in formulating scientific theories. What matters is whether mathematics plays a genuine explanatory role under a stable standard of adequacy.

Bibliography

Baker, Alan. (2022) “Simplicity,” Stanford Encyclopedia of Philosophy

Balaguer, Mark. (2024) “How to Be an Anti-Platonist,” in: Herbert Hrachovec and Jakub Mácha (Eds), Platonism: Proceedings of the 43rd International Wittgenstein Symposium, pp. 175-202 (Berlin/Boston: De Gruyter)

Balaguer, Mark. (2004) “Platonism in Metaphysics,” Stanford Encyclopedia of Philosophy

Daly, Chris and Simon Langford. (2009) “Mathematical Explanation and Indispensability Arguments,” Philosophical Quarterly 59 (237), pp. 641-658

Field, Hartry. (2016) Science Without Numbers: A Defence of Nominalism, 2nd ed. (Oxford: Oxford University Press)

Quine, W. V. (1948) “On What There Is,” Review of Metaphysics 2 (5), pp. 21-38

Sober, Elliott. (2015) Ockham’s Razors: A User’s Manual (Cambridge: Cambridge University Press)

EXPLAINING SMITH'S "STRIKING FACT": MOTIVATIONAL INTERNALISM VS. MOTIVATIONAL EXTERNALISM

Joseph Yoon

In *The Moral Problem*, Michael Smith observes that it is a "striking fact" that in the good and strong-willed person, a change in moral judgment is followed reliably by a change in moral motivation (Smith 1994, p.71). To explain this striking fact, Smith argues for motivational internalism, or the view that the moral judgment that it is right to do something is on its own sufficient to reliably produce moral motivation, and hence, morally virtuous behavior. He famously concludes that the opposing view, motivational externalism, or the view that something external to moral judgments, namely, a desire to do the right thing, is required to produce moral motivation, amounts to moral fetishism, and that being motivated by a desire to do the right thing is actually a moral vice (Smith 1994, p. 75). I will argue that Smith's arguments in favor of motivational internalism fail, and that far from being a moral vice, the desire to do the right thing is the one and only moral virtue, and it is the only thing that reliably ensures the moral motivation in an agent to do the right thing. The debate between motivational internalists and externalists is an issue of fundamental importance in ethics, as it asks what the source of morally virtuous behavior is: rational deliberation or desire?

Although I will mainly be focusing on Smith's argument for motivational internalism in this paper, it is helpful to consider what Smith's overarching project is. Smith's main task in *The Moral Problem* is not so much to argue in favor of motivational internalism as it is to solve what he dubs (surprise!) "The Moral Problem". The Moral Problem, as Smith formulates it, is the inconsistency of three statements that when taken each on their own, turn out to be plausible, but when taken together, turn out to conflict with each other. These three statements are:

1. Moral judgments express a subject's beliefs about an objective matter of fact, facts about what it is right for her to do.
2. If someone judges that it is right to x then *ceteris paribus*, she is motivated to x.

3. In order for an agent to be motivated to act in a certain way, she must have both an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences. (Smith 1994, p. 12)

The inconsistency can be detailed as follows: (1) tells us that the state expressed by a moral judgment is a belief, which (2) tells us it is necessarily connected to motivation in some way, and (3) that motivation being 'having a desire'. So (1), (2), and (3) taken together stipulates that there is some necessary connection between moral belief and desire, but (3) tells us that there is no such connection, as belief and desire are distinct existences (Smith 1994, p. 12). Smith aims to solve this problem by arguing in favor of motivational internalism. Although not the main aim of this paper, in the course of my argument, I will show that, along with his argument for motivational internalism, Smith's solution to the Moral Problem fails as well.

In section I, I will summarize Smith's argument for motivational internalism. In section II, I will give two *prima facie* arguments for why it fails. In section III, I will summarize weak motivational internalism. Then, I will show why weak motivational internalism fails in section IV. Section V will provide a summary of Smith's view of normative reasons. Section VI will show, against Smith, that normative reasons do not guarantee morally virtuous motivation. The next section, section VII, will detail two objections to reasons internalism and my responses. The conclusion will provide an argument for motivational externalism and concluding remarks.

I: Smith's Argument For Motivational Internalism And Humean Motivation

Smith first considers a version of motivational internalism which he dubs *strong motivational internalism*, which says that:

> If someone judges that it is right to x in circumstances C, then she is motivated to x in C.

This claim says that moral judgment brings about moral motivation *simpliciter*. Smith says, and I agree, that the claim is too strong because it commits us to denying that things like weakness of will may defeat an agent's moral motivations (Smith 1994, p. 61). He thus advocates for a modified version of motivational internalism which he calls *weak motivational internalism*, which states that:

> If someone judges that it is right to x in circumstances C, then either she is motivated to x in C or she is practically irrational (Smith 1994, p. 61).

This is the main claim that Smith seeks to defend in *The Moral Problem*, and it is the claim that I will argue against in the rest of this paper.

Before proceeding further, it is important to get clear on some terminology. Essential to either version of internalism is the concept of “right” (which Smith uses interchangeably with the term “moral”). What exactly is “rightness” or “morality” as Smith defines it? Smith writes, “...We can specify which property the property of being right is by reference to platitudes about rightness: that is, by reference to descriptions of the inferential and judgemental dispositions of those who have mastery of the term ‘rightness’” (Smith 1994, p. 39). He continues, "For example, there are platitudes concerning the *substance* of morality,” by which he means statements like: “Right acts are often concerned to promote or sustain or contribute in some way to human flourishing,” or “Right acts are in some way expressive of equal concern and respect” (Smith 1994, p. 40). Smith argues that it is such “distinctive *substance* or *content* of reasons that makes them into moral reasons as opposed to non-moral reasons” (Smith 1994, p. 183). From this, it seems that the definition of “right” that Smith has in mind is something akin to “morally virtuous,” which is the term I will be working with in this paper.

In order to argue for his view that moral judgments are sufficient to produce moral motivation, Smith appeals to the account of motivation provided by the Humean Theory of Motivation (HTM). Like Smith, I adopt this account of motivation as well, as I believe there are good arguments in favor of the view. Hume argues that the presence of both a *desire* and a *means-end belief* constitutes motivation, and moreover, that desire is necessary for motivation, because beliefs are mere representations of the world, and because they are not conative states in any way, are motivationally inert. The famous concept of “direction of fit” conveys this idea neatly. Beliefs are meant to fit with the world. They aim to be true, and inasmuch as they fail to fit with the world, they are false and should be discarded on those grounds (Smith 1994, p. 112). Desires move in the opposite direction. Desires are not meant to fit with the world, but rather, the world is meant to fit with our desires, at least in the eyes of the agent who possesses them. In general, agents mean to change the world to fit with their desires (Smith 1994, p. 112). Hume observes that desires and beliefs are distinct psychological states ("existences," in his terminology), because beliefs play a representational role, representing the way the world is, while desires play a motivational role, giving us the motivation to change the world to fit our desires (Radcliffe 2012, p. 121). These arguments, in my view, are strong arguments in favor of HTM. In any case, it is the theory of motivation Smith adopts, and therefore, the one I will be engaging with.

II: Why Motivational Internalism Fails

The main question then is: are reasons relative to desires, or are desires relative to reasons? That is, do certain facts about the world serve as normative reasons, which reliably produce certain desires in us and motivate certain actions on our part, or do our desires determine which facts count as reasons in favor of our acting a certain way (Smith 1994, p. 165)? Smith clearly wants to argue that desires are relative to reasons since his aim is to show that moral judgments, which in his view are rational cognitions, are sufficient to produce motivation on their own. He argues for this point by claiming that our beliefs about what it is right for us to do generate certain desires in us, and subsequently produce the Humean belief and desire pair that is required to produce moral motivation, and with it, morally virtuous behavior. For him, moral judgment, not desire, is the source from which morally virtuous behavior ultimately issues.

But, at this point, there are two ways to interpret Smith's argument for motivational internalism. One is an uncharitable view, and the other is a more charitable view, which although more charitable, reveals a much more fundamental problem with Smith's argument. I will start with the uncharitable view: If the foregoing is his argument for motivational internalism, it is fatally flawed, for motivational internalism is the argument that moral judgments *on their own*, without the additional contribution of any other factors, such as desires, are sufficient to motivate. Indeed, to argue otherwise would be to suggest that something other than moral judgment, or said another way, something *external* to moral judgment, is necessary to produce moral motivation, which is just motivational externalism. And indeed, Smith argues against this view. He is so opposed to externalism that he argues that not only is no factor beyond moral judgment itself required for motivation to occur, but that motivation that springs from anything other than moral judgments, such as a desire to do the right thing, amounts to a moral vice (Smith 1994, p. 75). But even if we concede that moral judgments are what initiates the causal chain that leads to moral motivation and morally virtuous behavior, Smith does acknowledge that desire, even where it is derived from moral judgment, is necessary for moral motivation. He writes, "All actions are indeed produced by desires, just as the Humean says; no actions are produced by beliefs alone…" (Smith 1994, p. 179). But if Smith concedes this, he is conceding that moral judgments, or our beliefs about what we have normative reason to do, cannot produce motivation on their own, and require something external to moral judgment, in this case desire, to produce motivation, and this is just the argument for motivational externalism. Thus, his argument for motivational internalism fails. Furthermore, any argument for motivational internalism that appeals to HTM, with its dictum that desire is necessary for motivation, is doomed to failure, due to its appeal to something *external* to moral judgment for the production of moral motivation. For that matter, any argument for motivational internalism that appeals to *any* factor external to moral judgment fails, because the motivational internalist has committed himself to the position that moral judgments are sufficient *on their own* to produce moral motivation. This also entails that Smith's solution to the moral problem fails, because the analysis of motivational internalism given above shows that motivational internalism can't be made consistent with Humean motivation, where desire and belief are distinct existences, and desire is necessary for motivation.

On the more charitable view, Smith's argument for motivational internalism is nowhere near as hostile to the role that desire may play. It makes room for desire to be a part of the story of motivation, as long as that desire is derived from rational judgment, or said another way, as long as that desire is relative to our reasons. But, in fact, Smith goes even farther than this and argues that desires are necessarily relative to our reasons, that they are always the product of our reasons. He argues that normative reasons, or what we are rationally justified in doing, are non-relative (vis-a-vis desire or anything else), but absolute (Smith 1994, p. 168). If this is true, then desire would necessarily have to be relative to reasons, since reasons are absolute and cannot be relative to anything else. And, if, as Smith argues, deliberating under conditions of full rationality always leads us to desire what is right, then reasons should always produce the morally right desires in us (Smith 1994, p. 184). But this view fails, because as Smith himself notes, desires and reasons can and do come apart, as in, we can desire to act in ways that contravene what we ourselves acknowledge that we have rational justification to do, and in such cases, it is desire which determines what we ultimately end up doing. Hence, it cannot be that reasons are absolute, or that they take priority over and above desires. Smith cites a number of examples given by Bernard Williams: "Consider the case of a woman who has a sudden urge to drown her bawling child in the bath; or the case of a squash player who, while suffering an ignominious defeat, desires to smash his opponent in the face with the racquet" (Smith 1994, p. 134). Williams says, and Smith agrees, that it is just false to think that the mother thinks she is rationally justified in drowning her child, and it is utterly false that the tennis player thinks he is rationally justified in harming his opponent. "But they desire these things none the less. They desire them in spite of themselves" (Smith 1994, p. 134). In both cases, the agent is able to come to a correct moral judgment as to what they are and are not rationally justified in doing. Clearly, then, it is possible for desires and reasons to come apart. They can be independent of one another. Reasons in no way guarantee that we will end up desiring to do the right thing. And when reasons and desires do come apart in such ways, it is desire that determines what we end up doing.

How do we know this? I think the empirical reality of evil in our world provides good evidence for this view. Take the example of a murderer who attempts to cover up his crime, as murderers often do. The fact that they attempt to cover up their crime indicates that murderers are aware that what they did was wrong. In any case, I think, and I think most people would agree, that it would be wrong to assume that murderers are people who mistakenly, but earnestly, happened to believe that committing murder was actually a good deed, that they only tried to cover up their crime because they believed that *society* disapproved, while they themselves maintained the belief that what they did was actually a good thing. Most murderers seem to be well aware of the fact that what they did was wrong. This means that they possessed the same capability as the morally good agent to come to the correct determination through rational deliberation of what they were and were not rationally justified in doing. And yet they still committed murder, which demonstrates that their desires, in this case, perverse desires, won out where their rational judgment and desires came into conflict. I would extend this argument and say that it most likely explains a good chunk of the evil phenomena we observe in our world. If an agent's grasp of what he was rationally justified in doing was all that was needed to guarantee moral behavior, the case of rational deliberators who commit evil acts in spite of knowing which acts are rationally justified would be impossible. And yet, the above examples show that such cases are possible, which undermines Smith's claim that reason will always lead us to desire the right thing and, in so doing, reliably produce moral motivation. Subsequently, this then undermines his motivational internalism. Reasons are clearly in no way absolute, or said another way, desires are in no way necessarily relative to our reasons. So, even if it is true on our more charitable view that Smith believes that desires derived from rational judgment do not count as being external to that judgment, and thus, are acceptable in his account of motivational internalism, motivational internalism still fails, because Smith's version of it claims that rational deliberation under conditions of full rationality will always produce in us the desire to do the right, and the above examples indicate that this is false.

III: Weak Motivational Internalism

But we noted above that Smith is not simply arguing for motivational internalism *simpliciter*, he is arguing for a weak version of motivational internalism. The weak version of motivational internalism that Smith advocates states that:

> If someone judges that it is right to x in circumstances C, then either she is motivated to x in C or she is practically irrational.

Put another way, weak motivational internalism says that if we are practically rational, then we are motivated to x upon judging that it is right to x. A few points of clarification are in order here. If motivational internalism is true, how does one arrive at a judgment as to what is right? Smith suggests an answer: "My suggestion then, in schematic form at least, is that our x-ing in circumstances C is right if and only if we would desire that we x in C, if we were fully rational…" (Smith 1994, p. 184). So, according to Smith, the right thing to do is what we would desire to do if we were fully rational. Moreover, for Smith, moral judgments are judgments about objective matters of fact, making him a realist and a cognitivist about morality. Given this, his argument is that rational deliberation leads us to a determination of what is morally right, and because he is committed to motivational internalism, he is also committed to saying that rational deliberation is ultimately practical in its issue, that it necessarily leads us to desire what is right, and therefore that it reliably produces and is the source of moral motivation and morally virtuous behavior in an agent. He writes, "...what we have normative reason to do is what we would desire that we do if we were fully rational" (Smith 1994, p. 150). However, one may, at this point, think that this view of motivational internalism might fail for the same reason given in the previous section, namely, that full rationality or practical rationality would in no way necessarily produce the desire to do the right thing. I would agree with such a view, but now I will go one step further and show that not only does full rationality not produce morally virtuous motivation necessarily, but that it is fully compatible with having perverse desires. I will show that the two are in no way mutually exclusive, which would undermine Smith's claim that the right thing to do is what we would desire to do if we were fully rational, full rationality being a summary notion of practical rationality according to Smith. And if this claim is undermined, Smith's assertion that practical rationality entails that we would be motivated to x upon judging that it is right to x would also be undermined, and with it, weak motivational internalism.

IV: Full Rationality, Perverse Desires, And The Failure of Weak Motivational Internalism

What, then, does Smith exactly mean by "full rationality"? First, Smith argues that full rationality is just a summary notion of practical rationality (Smith 1994, pp. 155-156). Second, he largely borrows his definition for full rationality (and thus, practical reason) from Bernard Williams, with some modifications. Williams elucidates, and Smith largely adopts, three conditions for full rationality:

1) The agent must have no false beliefs.
2) The agent must have all relevant true beliefs.
3) The agent must deliberate correctly (Smith 1994, p. 156).

But all three conditions are consistent with morally perverse motivation and behavior. Evil people, defined as people with perverse desires, can be fully rational: (1) An evil person can have no false beliefs about the way the world actually is, be under no illusions about what the facts of his circumstances are, and still have perverse desires. For example, an agent who desires to murder his next-door neighbor can have no false beliefs about his intended victim's movements, schedules, and routines on the day of the murder, and still possess the desire to murder him. (2) An evil person can have all the relevant true beliefs about the way the world is and still have perverse goals and desires. Nothing about the way the world factually is precludes an agent from having perverse goals and desires. For example, the same agent can have all the relevant true beliefs about where his victim will be at a given time on the day of the murder and what he will be doing at that time, and none of this information prevents the would-be murderer from continuing to possess the desire to kill his intended victim. If anything, it will aid him in his ignoble task.

(3) An evil person can deliberate correctly. Smith offers us four ways in which a person might deliberate correctly (three borrowed from Williams, and one that he develops himself), which all turn out to be consistent with a person having perverse desires and goals. First is means-ends reasoning. Smith argues that correct deliberation can generate new desires on the condition that we reason in accordance with the means-ends principle, "for only so does a desire for an end turn into a desire for the means" (Smith 1994, p. 157). However, a person with the perverse desire to rob a bank to enrich himself can generate a new desire to bypass the bank's alarm system by reasoning in accordance with the means-ends principle that his end of successfully stealing from the bank will be best served by bypassing the bank's alarm system. The second definition of correct deliberation that Smith offers is that of practical reasoning, as proposed by Bernard Williams, where "practical reasoning…[leads] to the conclusion that one has reason to x because x-ing would be the most convenient, economical, pleasant, etc. way of satisfying some element in…[one's set of desires]..." (Smith 1994, p. 157). Again, this definition of correct deliberation is still consistent with a person having evil desires. Returning to the bank robbery example, practical reasoning could lead the would-be bank robber to conclude that he has a reason to find a way to bypass the bank's alarm system because that will be the most convenient, economical, and effective way to satisfy his desire to successfully rob the bank. The third claim is that imagination too counts as a form of correct deliberation if the agent comes to believe that "he has reason to promote some development" upon exercising his imagination about what it would be like if said development were to come about (Smith 1994, pp. 157-158). This too is consistent with morally reprehensible behavior. The would-be bank robber may exercise his imagination about retiring to a life of comfort in the Caribbean with the stolen money and come to believe that he has reason to promote said development, in this case, committing a bank robbery. The fourth and last definition Smith gives for correct deliberation is that of systematic justifiability. Smith says that "by far the most important way in which we create new and destroy old underived desire when we deliberate is by trying to find out whether our desires are *systematically justifiable*" (Smith 1994, pp. 158-59). He elaborates:

> Suppose we take a whole host of desires we have for specific and general things; desires which are not in fact derived from any desire that we have for something more general. We can ask ourselves whether we wouldn't get a more systematically justifiable set of desires by adding to this whole host of specific and general desires another general desire, or a more general desire still, a desire that, in turn, justifies and explains the more specific desires that we have. And the answer might be that we would. For in so far as the new set of desires…exhibits more in the way of, say, unity, we may properly think that the new imaginary set of desires is rationally preferable to the old. For we may properly regard the unity of a set of desires as a virtue; a virtue that in turn makes for the rationality of the set as a whole (Smith 1994, p. 159).

But here again, this last definition of correct deliberation is still consistent with a person having perverse desires and ends. An evil person who has a specific desire to murder his next-door neighbor can obtain a more systematically justifiable set of desires by adding to his specific desire a more general desire to inflict pain and suffering on humans writ large, which in turn justifies and explains the more specific desire to murder his next-door neighbor. This set of desires would also exhibit more in the way of unity, and this unity would in turn make for the rationality of the set as a whole, albeit in a perverse way.

So, one can be fully rational and still have perverse ends. The two are completely compatible. Full rationality, and by extension practical reason, acts as no check, does absolutely nothing to preclude an agent from doing, or being motivated to do, morally reprehensible things. Conversely, it does absolutely nothing to reliably guarantee in an agent the motivation to do what is morally right, as Smith claims. The above examples demonstrate as much. Evil people, rational thought, and rational faculties are not at all mutually exclusive. Evil people can function just as rationally as you or I can. They are capable of having no false beliefs, possessing all the relevant true beliefs, deliberating correctly, and yet still choose to do morally reprehensible things. Rationality, it turns out, is neutral on matters of good and evil, virtue and vice. Rational thinking can be employed to realize virtuous ends just as effectively as it can be applied to realize perverse ends. Premeditated murder, especially one that has been meticulously planned weeks or months in advance, would be a good example of this.

A supporter of Smith's might object at this point that if moral realism is true, then full rationality, particularly the requirement that the agent possess all relevant true beliefs, would entail that a person would know what the right thing to do is in their circumstance, and this, combined with motivational internalism, would automatically motivate the agent to do the right thing. But still, this would not reliably guarantee moral motivation and behavior. As Smith himself points out, by way of Williams, our beliefs and desires can come apart. Let us return to the examples mentioned above of "a woman who has a sudden urge to drown her bawling child in the bath; or the case of a squash player who, while suffering an ignominious defeat, desires to smash his opponent in the face with the racquet" (Smith 1994, p. 134). As noted above, both Williams and Smith agree that it is false to think that the mother in question thinks she is rationally justified in drowning her child, and it is utterly false that the tennis player thinks he is rationally justified in harming his opponent. In both cases, the agents arrived at the correct moral conclusion as to what they are and are not rationally justified in doing. But again, Williams notes, and Smith agrees, "They desire these things none the less. They desire them in spite of themselves" (Smith 1994, p. 134). Clearly, then, one can be fully aware of what they are rationally justified in doing, which in this case would be that which is right, and still desire to act otherwise because they are unmoved by this information or they possess desires that are in direct opposition to it, e.g. they have a desire to do morally reprehensible things. These people do not suffer from practical irrationality as defined by Smith, since they possess no false beliefs, possess all relevant true beliefs, and can deliberate correctly. And yet they will still act in contravention of what they themselves believe they are rationally justified in doing because they possess a desire to do evil things. As shown above, in such situations, it is clearly our desires, perverse or otherwise, and not rational deliberation, that determines what an agent will end up doing. Hence, it is clear from the foregoing discussion that full rationality and practical reason in no way reliably guarantees that an agent will be motivated to do what is right, thus undermining weak motivational internalism as well. This, however, raises the question: what, then, is the correct account of the relationship between our desires and rational deliberation? Which account of this relationship most accurately captures the reality of how we make choices, and tells us which factor reliably determines and explains which choices agents end up selecting in moral situations? Smith offers us his theory with an account of normative reasons, to which I will now turn.

V: Normative Reasons

Smith begins his discussion by distinguishing two ways an agent might explain, or give reasons, for his actions: motivating reasons and normative reasons. Motivating reasons, according to Smith, give teleological, possibly even causal, explanations for an agent's actions by "citing the complex of psychological states that produce the action" (Smith 1994, p. 131). Normative reasons, on the other hand, "explain an intentional action in terms of the pattern of rational deliberation that either did, or could have produced it" (Smith 1994, p. 131). Such reasons constitute an agent's rational justifications for his actions.

This brings us back to the running thread of this paper: whether reasons, in this case, normative reasons, are relative to our desires, or if our desires are relative to our normative reasons. Let us recall what Smith's project is here. He wants to establish that moral judgments on their own are sufficient to reliably produce the moral motivation in an agent to do the right thing. Because he believes moral judgments are judgments about objective matters of fact, and therefore acts of rational cognition, he has to establish that doing the right thing in circumstance C is a necessary product of rationality, the thing we would do if we were fully rational. And indeed, he argues, "My suggestion then, in schematic form at least, is that our x-ing in circumstances C is right if and only if we would desire that we x in C, *if we were fully rational* [emphasis added] …" (Smith, 184). We have shown above that this is false, however. Full rationality in no way reliably leads a person to be motivated to do what is right but let us put this argument to the side for the moment. If normative reasons are reasons that "explain an intentional action in terms of the pattern of rational deliberation that either did, or could have produced it," and if we are deliberating under conditions of full rationality, then normative reasons would have to be practical in their issue in the same way that full rationality is (Smith 1994, p. 131). They too must reliably lead an agent to be motivated to do the right thing. And indeed, Smith writes, "...What we have normative reason to do is what we would desire that we do if we were fully rational," or put another way, what we have normative reason to do is to do the right thing, and since he is claiming that normative reason is practical in its issue, normative reason should reliably produce in us the moral motivation to do what is right (Smith 1994, p. 150). Moreover, if moral realism is true, and it is both full rationality and normative reasons that are the ultimate sources from which morally virtuous motivation reliably flows, then normative reasons must be absolute. If they aren't, then normative reasons can be a source of morally wrong behavior, and would no longer reliably serve as the source of moral motivation and behavior that Smith claims them to be, thus undermining Smith's claim that what we have normative reason to do are the things that are right, and suggesting that we could possibly have normative reasons to do evil. And indeed, Smith's position is that normative reasons are non-relative. The view of normative reasons that Smith espouses is called reasons externalism, which is the view that an agent's normative reasons for action are independent of her desires or subjective motivational set. I disagree with this view. I believe that reasons internalism, or the view that what we have reason to do is given by the content of our desires, offers a better explanation of human moral behavior (Smith 1994, p. 165). On this view, what counts as an agent's normative reasons for action are relative to and contingent on what the agent's desires or subjective motivational set happens to be.

But before I make the argument for reasons internalism, let us first analyze Smith's position in a bit more detail. Smith says, "What we have normative reason to do is what we would desire that we do if we were fully rational," and also, "our x-ing in circumstances C is right if and only if we would desire that we x in C, if we were fully rational" (Smith 1994, p. 131, 184). Hence, it follows that what we have a normative reason to do is what is right. In fact, Smith is claiming that we only have normative reasons to do things that are morally right, and not things that are morally wrong. Smith makes this clear in his discussion of the objectivity of normative reasons. He argues that because rightness is an objective matter of fact, it is discoverable through rational deliberation. Subsequently, he claims that everyone's desires would converge onto the same desire (the desire to do the right thing) if all parties involved in a debate on moral matters deliberate under conditions of full rationality (Smith 1994, p. 198). So he is claiming that given deliberation under full rationality, insofar as it necessarily leads us to the discovery of objective matters of fact (and given Smith's moral realism and moral cognitivism, morality is one of those discoverable matters of objective fact), *we will necessarily be led by full rationality to desire what is right* (we have shown in previous sections however that such a claim is untenable). If we take this along with Smith's assertion that "what we have normative reason to do is what we would desire that we do if we were fully rational," it follows that a normative reason is a reason to do what is right, not what is wrong, and necessarily so. Why does Smith go to such lengths to establish that normative reasons are absolute, and are the only reasons to do what is right and not what is wrong? Because he wants to bolster his claim that rational deliberation is the source from which morally virtuous motivation and action issues, and his view that normative reasons are a product of rational deliberation and are categorical (more on this point later), would mean that all rational people would do what they have normative reasons to do, thus ensuring that rational people, *by virtue of their reason*, will end up choosing to do what is morally right.

VI: Normative Reasons Do Not Guarantee Morally Virtuous Motivation

The idea that we only have normative reasons to do morally praiseworthy things and do not have normative reasons to do morally reprehensible things is, however, highly dubious. If we go back and look at what the definition of a normative reason is, we see that normative reasons "explain an intentional action in terms of the pattern of rational deliberation that either did, or could have produced it" (Smith 1994, p. 131). Returning to the bank robbery example, we could say that the bank robber has a normative reason to bypass the bank's alarm system, given his goal of stealing money from the bank. Rational deliberation would lead him to conclude that he would have a much higher chance of success if he disables the alarms, so his intentional act of disabling the alarms is explained by this bit of rational deliberation. It is absurd to claim that the bank robber does not have a normative reason to bypass the alarms simply by virtue of the fact that doing so is morally reprehensible. Put another way, it is absurd to claim that this bit of rational deliberation does not explain his intentional act of bypassing the alarms simply due to the fact that bypassing the bank's alarms is a morally reprehensible thing to do. Being morally virtuous is not a condition on whether something counts as a normative reason. In order for something to be a normative reason, all it needs to do is explain an intentional action in terms of the rational deliberation that produced that action, and rational deliberation is fully capable of explaining morally reprehensible actions as well as morally virtuous ones. It is absurd to say that a given normative reason does not explain an action, and therefore doesn't count as a normative reason, simply due to the fact that it is morally reprehensible.

In a previous section, we examined whether reasons broadly construed are relative to desires, or vice-versa. Let us consider now whether *normative* reasons are relative to desires or not. I have already mentioned that Smith's position is that normative reasons are non-relative and absolute. In fact, Smith is committed to making this argument, because if normative reasons are not absolute, that would entail that normative reasons do not *reliably* lead us to be motivated to do the morally right thing, which would undermine his entire argument. But how credible is the claim that normative reasons are not relative? As argued above, rational deliberation does nothing to ensure that we will always end up doing what is morally right. Again, using Smith's own definition of a normative reason, all something needs to do to count as a normative reason is to explain an intentional action in terms of the rational deliberation that produced it. If so, then any fact that explains an intentional action by providing a means-ends reason for pursuing a given course of action in light of what the agent's desires are, i.e. *relative* to an agent's desires, would count as a normative reason. And indeed, I think this view of normative reasons is a much more plausible one than the one Smith offers, because, as I have argued above, it is absurd to claim that a given normative reason fails to explain an action, and thus, fails to be a normative reason, simply by virtue of the fact that it is not morally praiseworthy. What we have reason to do is given by the content of our desires, and these desires can be perverse as well as noble, and the facts that explain these actions, whether noble or perverse, are normative reasons. Hence, it is true that the bank robber's desire to steal from the bank gives him a normative reason to bypass the bank's alarm system, that the would-be murderer's desire to kill gives him a normative reason to go out and buy a weapon. In both cases, normative reasons are relative to desires. And they do a perfectly good job of explaining intentional actions in light of the desires that produced them. In any case, it is absurd to claim that there are no explanations of evil acts because they are not morally virtuous. Evil acts have explanations, too, and whatever explains them are normative reasons. Therefore, clearly, there can be normative reasons to do evil things as well as good things. Whatever set of facts count as normative reasons are thus relative to what the agent's desires are. This, then, undermines Smith's view that normative reasons are absolute and only provide reasons to do what is morally right. Normative reasons are relative to our desires and could just as well give the rational agent a reason to do morally reprehensible things. This, by extension, also undermines Smith's overarching motivational internalist claim that rational deliberation reliably produces virtuous moral motivation in an agent.

Hence, I believe reasons internalism, where reasons are relative to desires, as opposed to Smith's morality-centric reasons externalism, is a more plausible explanation of human behavior. There is another major reason why Smith's position that normative reasons to do the right thing reliably produce moral motivation is highly dubious. Smith claims that normative reasons are categorical imperatives that apply equally to everyone regardless of what their individual desires may be (Smith 1994, p. 175). But if this claim is true, and it is also true that normative reasons are practical in their issue in that they reliably produce the motivation to do the right thing, then the mere existence of normative reasons to do the right thing would be enough to reliably guarantee moral motivation and action on the part of all agents, and we would be living in a perfect world (Smith 1994, p. 175). But obviously, we do not live in a perfect world, and I don't think Smith or anyone else is trying to mount a serious argument that we do. Evil is a part of the empirical reality of the world we live in, and Smith's view of normative reasons cannot explain this fact, because his view is that normative reasons to do the right thing are both categorical imperatives and are practical in their issue, which then entails that normative reasons must *reliably* produce moral motivation in *all* agents, owing to their categorical nature. Empirical reality emphatically refutes such a view, however. Thus, this is yet another reason why I think reasons internalism is a much more plausible explanation of empirical reality than the non-relative normative reasons version of externalism that Smith espouses. Reasons internalism is flexible enough to accommodate the very real presence of evil in our world, whereas Smith's version of reasons externalism is not.

VII: Two Objections To Reasons Internalism And Responses

Opponents of reasons internalism generally mount two objections to it, however. First, they find it unsettling that under reasons internalism, it is possible for an agent to have normative reasons to do morally reprehensible things (Asarnow 2016, p. 153). Second, they argue that reasons internalism suggests that there is no absolute morality, and that it suggests that there exists no objective normative reasons to act morally that are incumbent on all agents irrespective of what their individual desires and preferences may be (Asarnow 2019, pp. 34-35).

To the first objection, I don't see why this should be unsettling. Normative reasons themselves are neutral on matters of morality. All they do is explain an intentional action in terms of the rational deliberation that produced it. If a set of facts explains an act of evil in terms of rational means-ends deliberation that an agent conducted in light of his perverse desires, then those facts serve as normative reasons. Some philosophers seem committed to the idea that we should only have normative reasons to do the right thing, but the definition of normative reasons in no way requires this. As mentioned above, evil acts have explanations too, and if a set of reasons explain these morally reprehensible acts, they too count as normative reasons.

The second objection argues that reasons internalism suggests that there is no absolute morality, in that it makes reasons relative to our desires, and a person's desires may be noble or perverse, it implies that there are no objective normative reasons for moral behavior that apply to everyone, irrespective of individual desires. Smith himself argues against this view by asserting that normative reasons are categorical imperatives, not hypothetical (Smith 1994, p. 175). As mentioned above, however, the problem with this view is that it entails that normative reasons are enough to guarantee that everyone would be motivated to do the right thing, and it utterly fails to explain the very real phenomenon of evil in our world.

Furthermore, let us grant Smith and his colleagues this point. Let us assume for the sake of argument that there are indeed, right here and now, objective normative reasons for moral behavior that apply to everyone. It is still true that people who have such normative reasons may desire to act in a different way, and thus act in ways that contravene morality, no matter how categorical these normative reasons may be, because their desires come apart from their reasons. I cited two such cases in a previous section, as in the case of the woman who has an urge to drown her baby in spite of understanding that she has no rational justification, or normative reason, to act this way, and the case of the squash player who wishes to smash his opponent with his racquet, despite reaching the conclusion that he has no rational justification, or normative reason, to do so. Normative reasons and desires clearly can come apart, and when they do, it is desire that determines what we end up doing. If we assume that everyone has a normative reason to act morally, then even people with perverse desires have a reason to be moral. And this may well be true, but the reality of evil in our world demonstrates that people with perverse desires choose to and are capable of acting in contravention of these reasons, even if they really exist, and even if they are absolute, thus producing the evil that we all observe day to day in our world. Normative reasons are in no way capable of ensuring that agents will reliably be motivated to do the right things. And if we accept the Humean Theory of Motivation, this result should be unsurprising, as reasons are motivationally inert, and do not have the power to motivate on their own, much less reliably so. Thus, it is clear that no matter how categorical one may claim normative reasons are, and no matter how true such claims may be, normative reasons clearly do not have the capability to reliably produce morally virtuous motivation and behavior in us. Rather, it is desire that ultimately decides what we end up doing, whether in accordance with or in contravention of universal normative reasons of morality. Hence, I stress again that reasons internalism is a much more plausible account of human action than the theory that Smith proposes.

CONCLUSION

But an opponent of reasons internalism may point out that desires, as opposed to rational deliberation or normative reasons, are no more reliable than these in producing moral motivation and action. With this I would agree. Then what factor, if any, can *reliably* produce moral motivation and behavior in agents? If reasons, normative or otherwise, categorical or not, cannot reliably guarantee moral motivation, nor can full rationality or rational deliberation, then what can reliably produce in agents the motivation to do the right thing, especially in light of the fact that reasons are relative to desires, perverse or not? The desire *to do the right thing* can. An agent in possession of such a desire would have the desire and goal to do what is morally right, and this would supply him with normative reasons to do the right thing. A desire *to do the right thing* would ensure that an agent possessed of such a desire will *reliably* be motivated to do what is right, irrespective of the facts or circumstances presented to him. And because he has this desire, it would also ensure that all his reasons for action are reasons to do the right thing. Even if we embrace the Humean Theory of Motivation, this view will still hold, in that desire and beliefs can remain separate entities, and yet, still produce the motivation in an agent to do the right thing. The necessary component, desire, would be aligned with right actions, and means-ends beliefs about how best to achieve said desire would reliably produce the motivation in an agent to do the right thing. Hence, far from being a moral vice, the desire to do the right thing is the one and only moral virtue, the only thing that ensures that an agent will be reliably motivated to do the right thing, and motivational externalism is a better explanation of the "striking fact" than motivational internalism.

Bibliography

Asarnow, Samuel. (2019) “Internal Reasons and the Boy Who Cried Wolf,” *Ethics* 130, pp. 32-58.

Asarnow, Samuel. (2016) “Rational Internalism,” *Ethics* 127, pp. 147-178.

Radcliffe, Elizabeth S. (2012) “The Inertness of Reason and Hume’s Legacy,” *Canadian Journal of Philosophy* 42, pp. 117-133.

Smith, Michael. (1994) *The Moral Problem* (Malden: Blackwell Publishing).

Boys Have Shown Their Will to Change: A Philosophical Account of Boyhood, Emotion, and Becoming in Novel

Rafi Bustos

Introduction

Patriarchal social arrangements are ubiquitous in the history of philosophy and can be traced back to the ancient philosophers such as Aristotle who described the nature of woman as a "deformed man" (Aristotle 2014, pp. 1131-1133). Common discussions in feminist philosophy have usually concerned themselves with how patriarchy perpetually oppresses and harms women. Yet much less has been discussed about its effects on men, and more specifically, boys. By "patriarchy" I mean something similar to what the feminist theorist and social critic bell hooks means when she describes patriarchy in her 2004 book *The Will to Change* as a system "that insists males are inherently dominating, superior to everything and everyone deemed weak, especially females" and that maintains this dominance through "psychological terrorism and violence" (hooks 2004, p. 18). Most assume that under patriarchy men (and boys) "win", as they are seen as superior and dominant to women. However, one ought to ask, as hooks does, "Does patriarchy *accept and celebrate* men and boys? Or are they too, forced into rigid, even perhaps violent and dangerous, gender roles like women and girls?" She addresses this, writing that one of the "tremendous failings" of feminist theory and practice is the lack of studies on boyhood or any alternatives to counter patriarchal masculinity (ibid., 39). She also discusses that one reason countering patriarchy is difficult is because most are "fixated" on furthering patriarchal attitudes through children's, particularly boys', literature and media (ibid., 53). For hooks, masculine media perpetuates and pushes the social belief that boys and men's emotional identities are to be "inherently" violent and emotionally inept beings.

These days, it is hard to disagree with hooks' assertions, given the popularity of hyper conservative and misogynistic "red pill" media grabbing the attention of many impressionable young boys. [1] However, I argue that she heavily undermines media's (and by extension men and boys') capacities of *challenging* patriarchal messages. Instead, it has *already* given us alternative (i.e. non-patriarchal) forms of masculinity. Olufemi Taiwo in *Stoicism (as Emotional Compression) is Emotional Labor*, presents a nuanced commentary of hooks' argument by stating men could have a form of "feminist" and "prosocial" masculinity in emotional work through the practice of "emotional compression" which he believes is the "more accurate" form of philosophical stoicism (Taiwo 2020). My paper does not decide whether hooks or Taiwo's commentary on masculinity is more accurate. Rather, I attempt to reconcile the two arguments by considering a philosophical perspective of emotions, mainly anger, according to Elizabeth Spelman's logic of anger (1986). I shall then propose an alternative to patriarchal masculinity we, as feminist philosophers and thinkers, must strive for. This possible alternative equates masculinity to Helene Cixous' notions of "feminine and philosophical becoming" in art and philosophy with support of feminist writer, Audre Lorde. I will present my argument through two coming-of-age novels: *The Outsiders* by S.E. Hinton (1967) and *The Body* by Stephen King (1982).[2] In the first part, I will discuss *The Outsiders*, and how its story depicts the consequences of unchecked patriarchal violence and anger for boys. Secondly, I discuss *The Body* and how its characters and themes directly challenge patriarchal conventions through the emotional labor Taiwo suggests. Finally, I shall use what I analyzed in both works to argue for and offer a masculinity that entails a philosophical becoming reminiscent of Cixous and Lorde's aesthetic views on philosophy and emotions respectively.

PART I: *THE OUTSIDERS* BY S.E. HINTON

The Outsiders by Susan Eloise (S.E.) Hinton is about a group of lower-class Oklahoman boys, primarily known as "Greasers," who are constantly in conflict with the upper middle-class gang known as "Socs" (short for "Socials"). The story's protagonist, Ponyboy Curtis, a fourteen-year-old Greaser, accidentally becomes an accomplice to a murder after his friend, Johnny, kills a Soc in self-defense.

Although most readers can probably deduce the sociocultural commentary of the story (Beals 2018, p. 183), there is also an account of masculinity and emotions, both the kinds that are encouraged and repressed, developed throughout. By the end of the novel, Ponyboy not only sees both groups the same, namely as victims of a society that has forced them to bear the consequences of a highly normalized patriarchal and violent culture. He comes to this conclusion through his perception of emotions.

1.1 FEELING IN THE OUTSIDERS

As bell hooks writes, “Patriarchal boys, like their adult counterparts, know the rules: they know they must not express feelings with the exception of anger” (hooks 2004, p.42). She describes this as “male stoicism.” Her commentary is reflected in Ponyboy’s conversation with Cherry Valance, a Soc girl he befriends at the movies. As the two converse and are surprised to enjoy each other’s company, they discuss what they believe are the differences between Greasers and Socs:

> “No,” Cherry said slowly when I said this, “It’s not just money. Part of it is, but not all. You greasers have a different set of values. You’re more emotional. We’re sophisticated—cool to the point of not feeling anything. Nothing is real with us…”
>
> […] “That’s why we’re separated,” I [Ponyboy] said, “It’s not money, it’s feeling—you don’t feel anything and we feel too violently.”
>
> “And”—she was trying to hide a smile— “that’s probably why we take turns getting our names in the paper” (Hinton 2006, pp. 30-31).

This excerpt reveals precisely the type of emotional masculinity that, as hooks suggests, patriarchy enforces. Cherry describes how Socs and Greasers not only differ economically, but also in how they “feel.” For Socs, who feel *too little*, become, as hooks describes, stoic. On the other hand, Greasers feel *too much*, which manifests as violence. While readers may take away from this excerpt that this conversation implies a sort of “natural” difference between the groups’ ways of “feeling,” while simultaneously denying a social factor (i.e. money) for their differences, I argue that these differences in “feeling” and social factors are not mutually exclusive. Indeed, Cherry insists that money is not the *direct* reason for the difference between the Greaser and Socs’ emotional expressions. However, I believe it certainly has an *indirect* influence on them. In her 2021 article on *The Outsiders,* Swedish Young-adult fiction scholar Lydia Wistisen, supports this identification and distinction between the two gangs’ emotions by suggesting the Socs make emotional suppression “cool” in an attempt to express their socioeconomical superiority in terms of “maturity” and “successfulness”—a reflection of a post-war United States culture in which most middle-class men exhibited “manhood” through emotional suppression and numbness. On the other hand, “the delinquent” (or the Greaser), is seen as someone who desires “individual autonomy, power, and control” in opposition to these middle-class masculine expectations (Wistisen 2021, pp. 205-206). Regardless of their differences, she believes that *both* forms of emotional expression and management prove to harm *both* groups of boys, as when Ponyboy and Cherry joke that this is what makes them "take turns making it on the paper" and getting into legal trouble (ibid.).

However, I would like to introduce a caveat to hooks' statement, particularly on the notion of "stoicism." The philosopher Olufemi Taiwo offers a more nuanced understanding of the stoicism that hooks condemns, distinguishing between "Ancient stoicism" and "modern stoicism" (the colloquially "appropriated" use of the term) (Taiwo 2020, p. 16). "Modern stoicism," *is not the same* as its ancient philosophical counterpart, which teaches emotional management through distinguishing what is "one's own" to "what is not"; (Epictetus 1904, p. 1-5). Modern stoicism "lacks self-knowledge of one's emotions" (i.e., alexithymia) and encourages emotional unavailability to care for oneself and others' emotions. He emphasizes that these two aspects of "modern stoicism," paired with "the toxic presence of homophobic and misogynist norms that complicate homosocial forms of support," result in the kind of masculinity prevalent in patriarchy (Taiwo 2020, p. 5). This is important, because another possible reading of this conversation between the Greasers and Socs illuminates these two features of the "modern stoic" that Taiwo describes. Although we can represent the Greasers as "angry" and Socs as "stoic," we can also represent them as exhibiting "alexithymia" and "emotional unavailability" respectively. As Taiwo explains, alexithymia involves a *lack* of knowledge about one's emotions. It is evident throughout the story (and in my later analysis in 1.2) that while Greasers are permitted to "feel," they do not have the *knowledge* to identify the "feeling" or the emotion. [3] On the other hand, the Socs' belief that they are "too cool to feel" anything is synonymous with a form of "emotional restriction" that involves "emotional unavailability," the suppression and inability to communicate emotion.

Wistisen's theories and Taiwo's philosophy allow us to see clearly an image of the problematic "modern masculinity" Hinton attempts to depict in her novel. It is evident in this small exchange between Ponyboy and Cherry that while there is a distinct economic disparity between Socs and Greasers, *both* respond to emotions similarly: through violence. While the Socs "feel nothing," they repress emotions until they manifest it into physical and verbal harassment towards the Greasers. On the other hand, while Greasers have fewer expectations from society to suppress their emotions, they are unaware of their emotions and recklessly express emotion (or "cool-off" as employed by the novel) through petty crime and turf wars with the Socs. The former's emotional suppression and the latter's alexithymia eventually devolve into violence towards each other. And with these commentaries in mind, I believe hooks' argument can potentially be seen as overgeneralizing due to her misuse of the language of "stoicism."

1.2: The Consequences of Patriarchy: Violence and Anger

In hooks' earlier claim, anger is an emotion heavily involved in the upbringing of patriarchal men. She also explains, "Much of the anger boys express is itself a response to the demand that they not show any other emotions. Anger feels better than numbness because it often leads to more instrumental action… things that are waiting to be heard" (hooks 2004, p. 44).

Here, hooks seem to understand anger in terms of a cognitivist theory of emotion. According to Elizabeth Spelman, a cognitivist theory depicts anger, not as an irrational urge or mere bodily symptom, but a form of judgment that contains a distinct "logic." Anger, for the cognitivist, is not directionless, but has a direct intended object. Someone who expresses anger believes they were wronged or given unjust treatment (Spelman 1986, p. 266). Sometimes, when anger expressed by subordinated groups goes unnoticed, it can devolve into violence. To mitigate devolving to violence, people must understand that subordinates have a "right" and even more, "ought" to express their anger to the dominant forces that often exploit them (ibid., p. 271). Audre Lorde shares a similar notion, stating that anger is a "well-stocked arsenal" that is "useful against oppression, personal and institutional" (Lorde 2012, p. 127).

All three writers seem to agree on what anger is: informative and a sign of injustice. Although Spelman specifically describes the anger of women, and for Lorde, black women, they *all* adopt a "cognitivist view" of anger, which states that anger responds to a perceived injustice, harm, or slight. And I argue that this view of anger reveals the consequences of patriarchy's insistence that boys sever ties with their emotions. *The Outsiders* illustrates this through the character of Greaser, Dallas "Dally" Winston.

Dallas, immediately from the first chapter, is described as "cold," "hard," and "bitter" (Hinton 2006, p.10). He grew up in the New York organized gang scene and was abused by his father as a child. The only way to "survive" in those conditions was to become hardened by his environment. Ponyboy consistently describes him as an "angry" and highly aggressive person throughout the novel. We see this when Dallas gives advice to Ponyboy, saying, "…You'd better wise up, Pony…you get tough like me and you don't get hurt. You look out for yourself and nothin' can touch you…" (ibid., 112). It is evident that Dally embodies the patriarchal masculine norms of toughness and hyper independence. To survive the violent patriarchal world that he was thrust into alone, at a young age, Dallas learned he must "play" by its rules.

However, Dally's character is complex. He shows affection and care towards his friends, as seen especially with his relationship with Johnny. When Johnny murders Bob, Dally immediately comes to their aid and helps plot their escape. However, despite all of his efforts to protect them from the police and even financially provide for the boys during their hiding, Johnny eventually dies after deciding to risk his life to save a group of children from a fire. This sends Dallas emotionally into turmoil. Dally's grief turns into a form of anger. After attempting to rob a grocery store and provoke the police with his unloaded gun, Dallas is shot to death. Ponyboy believes "this is what Dally would have wanted," implying he committed suicide (ibid., 117).

The tragedy of Dallas Winston can be traced back to his anger. Applying the cognitivist view, a possible interpretation of his (and many other boys') anger, identifies the object of Dallas's anger as the system that made him. He is angry at the environment and the way society has forced him into this situation. Although he abided by patriarchal norms and perpetuated it in his behavior, the system proceeded to take more from him through Johnny. And as explained earlier, Dally had never been taught how to identify or understand his emotions because of his violent and patriarchal upbringing. As a result, Dally died under the very system that he had been raised under. Per Spelman and Lorde, violence is evidence that reveals Dally's "anger" was never taken seriously.

I want to clarify that I am *not* defending the actions that Dallas' anger motivated. But as I established earlier, what other way *could* he have approached it when he, and many of the other Greaser boys, have not been taught the proper knowledge or skills to identify and understand their own emotions? Patriarchal thinking never gave them the resources to allow them to do so and merely taught them to "cool-off" using violence. I *do* however want to emphasize what the anger he felt *was*. As Spelman and Lorde would have affirmed, Dally has a "right", and in fact, "ought" to be angry. Society primed him to be violent, cold, and bitter. And whilst he obeyed and practiced its "laws," he *still* lost what little "non-patriarchal" aspect he had of him through Johnny's death. Ponyboy even remarks that had his parents been alive, they would have emotionally supported Dallas and helped him see "there is still good in the world" (ibid., 38). If Dally was taught a non-patriarchal way of handling his emotions, he would have possibly handled the anger and grief over his friends' death more effectively and evaded his tragic end. Therefore, when one sees anger in this manner, Dally's death depicts the direct consequences of teaching patriarchal masculinity for boys.

Furthermore, the story of *The Outsiders* not only condemns the patriarchy hooks believes the media tries to enforce and perpetuate but also exhibits the consequences boys face when having their entire boyhood dedicated to upholding patriarchal standards. Hinton shows that patriarchy prevents boys from realizing they are victims—they *all* are "losing" even if they are told they are "cool" or "tuff" for sacrificing their emotional lives. It creates a life forcefully lived in alexithymia and emotional unavailability which festers into anger. This anger, when ignored and uncared for, becomes self-destructive and violent. Ponyboy's older brother, Sodapop, observes this, saying, "…If you don't have anything, you end up like Dallas…and I don't mean dead, either. I mean like he was before. And that's worse than dead…" (ibid., 135).

PART II: THE BODY BY STEPHEN KING

While *The Outsiders* cautions its readers about the consequences of patriarchal masculinity, Stephen King's *The Body* portrays a possible alternative to it. The story revolves around a group of boys who attempt to travel through the woods of Maine to find the corpse of a missing boy. King's novel does not necessarily focus on the search for the dead boy, but rather on the interpersonal relationships and emotions shared between the boys on their journey. As I'll show, the way the characters express intimacy resonates with Taiwo's theory of emotional compression. As Taiwo argues, this form of emotional labor is potentially "prosocial" and "feminist," outlining a potentially *non*-patriarchal emotional landscape for contemporary masculinity (Taiwo 2020, p. 6).

2.1: EMOTIONAL COMPRESSION IN RELATIONSHIPS AND THE "ESCAPE"

Earlier, I discussed how Taiwo distinguishes "modern stoicism" from Ancient philosophical stoicism. In turn, he describes the latter's views on the emotions as the basis for his theory of "emotional compression." Taiwo emphasizes that emotional compression comprises "fullness" and ability to "communicate them [emotions] clearly and fully, while being guarded in the public *performance* of the emotions…Emotional compression done well involves tightly managing one's actions importantly including this aesthetic of character and its contribution to one's actions' expressive content" (ibid.). bell hooks supports a similar form of emotional management for men when she describes what “feminist masculinity” entails, writing, “…feminist masculinity defines strength as one’s capacity to be responsible for self and others” (hooks 2004, p. 117).

Based on Taiwo’s description of emotional compression and hooks’ description of “feminist masculinity,” I found the character of Chris Chambers to embody these aspects. He, alongside the protagonist, Gordie Lachance, is portrayed as the emotional anchor of their group. Throughout the story, he settles their disputes, comforts them, and ensures none of them are ostracized or left behind. One of his most iconic scenes in the novel (and film) was his argument with Gordie. Chris urges him to become a writer and take college preparatory courses regardless of what his family or friends do to “drag him down.” Gordie denies his own writing skills and believes that he would rather stay in the shop courses with his friends than be alone and ahead. Frustrated after Gordie remarks about how he acts like a father, Chris retorts:

> “I wish to fuck I *was* your father!” […] “You wouldn’t go around talking about takin those stupid shop courses if I was. It’s like God gave you something, all those stories you can make up, and He said: This is what we got for you, kid. Try not to lose it. But kids lose *everything* unless somebody looks out for them and if your folks are too fucked up to do it then maybe I ought to” (King 2018, pp. 108-109).

Even the narrator, Gordie, was taken aback by Chris’ remarks. It surprised him how eloquently Chris spoke yet was struck with “terror” by his words (ibid.). Chris was (understandably) frustrated over Gordie’s reluctance and self-sabotaging behavior. Based on the tone and diction, he is clearly angry. Yet, he can effectively and clearly share the reasons for his frustrations towards his friend without resorting to violence or hostility. He directly tells his friend his *exact* reasons for expressing his current feelings. Furthermore, he is willing to take emotional responsibility for *others* as well. Here, Chris successfully expressed emotional compression, and it proved to be effective due to how Gordie receives it. Gordie does understand his friend’s emotions and responds to them accordingly. This conversation between the two becomes vital to their character progression in the novel. Chris eventually persuades Gordie to be an author despite his family’s neglect. In turn, Gordie encourages Chris to realize his goal of leaving his own abusive family to improve himself.

This excerpt reveals the effectiveness of Taiwo's proposed emotional labor. The character was not excessive nor deficient in his emotional expression. He can clearly identify the emotion he experiences and, as mentioned earlier, is emotionally available for his friends as he takes on the roles of mediator in the group. Chris also states he is willing to take an emotional and "fatherly" responsibility for his friend since he knows Gordie's parents are "too fucked up" to care about his aspiration due to their extreme grief over their older son, Dennis. From this scene, we are shown what healthy friendships (and by extension healthy alternative masculinity) entail. It is the ability to not only identify and acknowledge one's *own* emotions, like Taiwo describes, but the ability to be emotionally available, communicate, and effectively manage the emotions of others. And this in turn makes his characterization unique.

One could argue as well that the boys' desire to "escape" Castle Rock is to "escape" the masculine expectations placed upon them by their families and their community. *The Body,* like *The Outsiders,* is set in postwar 1960s America wherein the expectation for most boys and men is dictated by their family, class, and society.[4] Gordon comes from a lower middle-class family. His family expects him, as a boy from lower middle-class, to be like his older brother who was a model "middle-class American boy" who was athletically gifted and served his country. But when he fails to be like Dennis, they become disinterested in him. Chris, whose family relies on benefits from the government (that his father uses to buy alcohol), is from the working-class. Everyone assumes he will eventually become a "delinquent" simply because of his familial ties.

These expectations in both the middle and working-class boys are teeming with patriarchal expectations. Patriarchy, regardless of class, seems to enforce on the boys that they are *only* valued if they can adhere to its definitions and standards of "manhood." If they are *not* compliant with them, then they are immediately punished and alienated. Gordie fails at becoming the "middle-class American man" his brother was and is reprimanded through emotional neglect. Chris, actively rejecting the assumptions and labels given to him as a working-class boy from a troubled home, is physically abused by his family and rejected by his community for wanting to change. Despite the alienation, abuse, and rejection, they chose to double down. Both boys eventually take the junior high college courses, finish high school, and attend college. While Gordie eventually becomes a teacher and successful writer, Chris goes to law school to become a lawyer. However, it is revealed that Chris dies after attempting to stop a fight at a restaurant wherein he was stabbed in the neck (ibid., 177). Out of all their friends, Gordie is the sole remaining person alive by the end of the story.[5]

The theme presented by *The Body* of "escaping" expectations on oneself not only applies to goals and ambitions, but also to expectations given by society, class, and culture. Particularly in this case, values and norms expected in men. While the protagonist successfully escapes the societal and masculine expectations placed on him, patriarchy proceeds to persistently reprimand him. His friends who desperately wanted to submit to social stereotypical male forms of expression in the 1960s, died gruesome deaths attempting to do so. His closest friend had his life taken away by a man who irrationally reacted with violence. The conclusion of Gordie being the sole survivor amongst his friends is poignantly symbolic. Although he was able to define himself on his own terms as both a writer and in how he performs manhood through his unique expression of emotion and his relationship with the late Chris Chambers, it came at the cost of complete ostracization from patriarchal society. *All* his friends met their demise at the hands of patriarchy (directly, with Chris; indirectly with Vern and Teddy). The same group of friends that Chris repeatedly warned Gordie would "drag him down."

The Body exemplifies an alternative form of masculinity hooks believe is difficult to find in media. Through the relationship and characters of Gordie Lachance and Chris Chambers, we are shown that Taiwo's solution of emotional compression is a possible and achievable form of masculine expression that we can provide to men to allow for more fulfilling and healthier emotional lives. Their ability to manage each other's emotions equitably and fairly through the act of emotional compression results in both boys ultimately reaching their goals. And although patriarchy continues to reject these boys' active resistance to its norms and even subjugating them to direct violent consequences, the boyhood Lachance and Chambers embodied in the novel prevailed. This story not only eloquently illustrates nonpatriarchal masculinity's potential but also shows that it *is* possible to "escape" the banal definitions of patriarchy. While the ending of the novel is bittersweet, it leaves its readers a with message that implies that the boys who chose to "escape" the grasp of patriarchy lived more meaningful lives than those who chose to abide by it.

Conclusion: The Philosophical Boyhood

Through the analysis of both novels and their depictions of boys' emotions and masculinity through a philosophical lens we see how literature has already presented different ways of counteracting patriarchal culture. *The Outsiders* criticizes the violent and emotionally depriving environment that becomes the setting of upbringing for most boys and its consequences. *The Body* shows boys' defiance and rejection of patriarchal expectations imposed on them and their eventual successful (albeit melancholic) escape from it to pursue more creative, emotional, and meaningful lives. With these, I want to attempt to answer the question bell hooks and many other feminist philosophers pose: If patriarchal masculinity is failing our boys and men, what kind of "model" (if there is any) of masculinity ought we teach them? The answer to me would be "a philosophical one." What does it mean for men to have a philosophical becoming?

To answer both questions, I will refer to Cixous' philosophical and literary approach to the notion of becoming in her 1987 paper, "Reaching the Point of Wheat, or a Portrait of the Artist as a Maturing Woman." She begins by prefacing that although she uses the term "feminine" and "sexual difference," throughout her paper, it does not mean that her writing and philosophy imply there is an "essential difference" between men and women. Rather, she uses this kind of language to pertain to a certain kind of "becoming" that has been associated with the depictions of women and femininity in literature.

For Cixous, the artist's education through "libidinal" structure begins with "taste." She uses the term "the primitive scene" from Freud's psychoanalysis to describe the "initial" process of attaining knowledge. She uses the story of Genesis in the bible where Eve is tempted into eating the apple, the fruit symbolizing knowledge, to exhibit this process:

> So the message is: "Don't." That's all: "Don't." And there is the other message, that of the apple, which says: "Try me, I am beautiful." There is no reason why she should not try, because the death message is meaningless. So she tries because she is a woman. That is what the bible says, and it is probably true.
>
> I think it is true that her decision must have been determined by something "feminine" in her structure, particularly her desire and her non-fear of *knowing what is inside.* So knowledge started for all of us knowing with the mouth, by tasting. Taste is the first act of knowledge, for women and for all men who are women. And the price of it has been exile, death, but also work, art, creation. (Cixous 1987, p. 3)

Cixous believes that "true" knowledge is shrouded behind prohibitions. Here, Eve, the woman, defies the restrictions put against her to attain knowledge. She does so by using "taste," the first stage in Freud's psychosexual development. People's first way of expressing pleasure *and* understanding the world is through the mouth. However, the pleasure of earning this "knowledge" comes at a cost. In the story, God exiles humans out of paradise as punishment for eating the forbidden fruit. Human's acquisition of knowledge out of pleasure usually involves the questioning and direct confrontation with "law." The "pleasure" of learning and creation of knowledge is restricted, forbidden, and shamed.

She emphasizes this theme of the artist's relationship to "law." While laws can be de jure, or be established in legal code, the "law" that Cixous frequently invokes throughout the essay is de facto law, which are normalized beliefs, values, and standards in society. She asserts using the works of James Joyce and Clarice Lispector that the knowledge acquired from pleasure for artists and "women" first comes in an act of "questioning" or "undermining" these de facto laws.[6] These characters reject the law because of an inner and deeper desire that is seen by society as shameful (hence for Cixous, she believes the law finds people "already guilty"). To her, the artist is a person who becomes "indifferent" to the law and can creatively find ways to express knowledge through pleasure.

I believe that Cixous' philosophy helps express the core of my notion of the philosophical boyhood. The questioning of "law" is perfectly exemplified in both *The Outsiders* and *The Body*. Both its main characters, Ponyboy and Gordie respectively, have a somewhat strained relationship to the "laws" of their society. One can read Ponyboy and Gordie to already be at odds with the law for being more "emotional" than the other boys and men in their lives. They are emotionally sensitive compared to the others in their stories. This emotional sensitivity we see in their character already conflicts with the patriarchal expectations for boys to be emotionally detached. But this also makes them narrators of a potential and "feminine" becoming of men. These two characters' direct conflict with the de facto masculine laws of their world is the catalyst of their entire "coming-of-age" thematic, their "becoming" into men (or as Cixous puts, men who *are* "women"). They *are* characters indifferent to the law. Ponyboy is from a gang of boys who are deemed social outcasts. Yet, he also begins to question the values of his *own* outcast group, particularly how they express masculinity through anger and violence. Gordie struggles to conform to expectations set for him by his family and community as they compare him to his late older brother. Both boys are wrestling with the laws of patriarchal masculinity imposed on them. Their journey throughout their stories is about finding "knowledge" (or perhaps "truth") of their situation in hopes of reclaiming the deep desire that patriarchy continues to shame and prevent them from feeling. And they do so out of a forbidden masculine desire: love.

Interestingly, we find a similar notion of the knowledge derived from "desire" and "pleasure" from Audre Lorde in her notion of *eros* in *The Uses of the Erotic*:

> The very word *erotic* comes from the Greek word *eros*—the personification of love in all its aspects—born of Chaos, and personifying creative power and harmony. When I speak of the erotic, then, I speak of it as an assertion of the lifeforce of women; of what creative energy empowered, the knowledge and use of which we are now reclaiming in our language, our history, our dancing, our loving, our work, our lives. (Lorde 2012, p. 89)

Lorde continues to clarify what she means by the term "erotic." She believes that men have reduced and "plasticized" the erotic to be associated with sexuality and pornography. Eroticism has a more nuanced, complex, and deeper meaning. Both Lorde and Cixous seem to aesthetically share this meaning *eros*, pleasure, as both a creative and desiring force. To them, it is also the desired love in its most unadulterated form. But as we have seen throughout, this "desire" has been slandered, demonized, and prohibited by patriarchy. It is absent from boyhood and continues to be absent throughout manhood. It is the core of patriarchal law to not only appropriate emotions as "feminine in nature," but to vilify emotional life for men and boys. The inability of men and boys to be in touch to their emotions and desiring love, results in angry men who frequently resort to violence. As hooks remarks:

> Patriarchy has sought to repress and tame erotic passion precisely because of its power to draw us into greater and greater communion with ourselves, with those we know most intimately, and with the stranger. Feminism changed the intimate lives of women and men by offering a vision of relationships rooted in mutuality, a vision of partnerships without domination. This seductive promise can be fulfilled only as patriarchal thinking ceases to dominate the consciousness of women and men, girls and boys. Seeking to heal the wounds inflicted by patriarchy, we have to go to the source. We have to look at males directly, eye

> to eye, and speak the truth that the time has come for males to have a revolution of values. We cannot turn our hearts away from boys and men, then ponder why the politics of war continues to shape our national policy and our intimate lives. (hooks 2007, pp. 183-184)

It is also thematically fitting to Cixous and Lorde's philosophy that both boys *become* artists by the end of their novels. Both boys become writers of their stories so that they may be able to teach others, other boys, that they are not alone in their struggles. By the end of *The Outsiders*, whilst processing the death of Johnny and Dally, Ponyboy explains why he chose to write this story:

> …Suddenly it wasn't personal. I could picture hundreds and hundreds of boys living on the wrong sides of cities, boys with black eyes who jumped at their own shadows. Hundreds of boys who maybe watched sunsets and looked at stars and ached for something better. I could see boys going down under street lights because they were mean and tough and hated the world, and it was too late to tell them that there was still good in it, and they wouldn't believe you if you did. It was too vast a problem to be just a personal thing. There should be some help, someone should tell them before it was too late. Someone should tell their side of the story, and maybe people would understand then and wouldn't be so quick to judge a boy by the amount of hair oil he wore. It was important to me. (Hinton 2006, p. 136)

For Ponyboy, writing his story is not only his way of showing his indifference to the law, but also his way of showing other boys in the same position as him that it *is* possible to live *beyond* the reach of patriarchal society. It is possible to become a man without the need to resort to violence. He comforts these boys, showing them that their frustrations and anger towards the world are reasonable. Most importantly, he hopes that in writing his experience, in becoming an artist, he can warn others about and prevent what happened to Johnny and Dallas from happening to other men and boys. Ponyboy's deep love for his friends and his ability to openly express this love—this *eros*—in writing allows him to become a man-who-is-woman. He validates the anger, pain, and hurt that boys experience under patriarchy. Their anger *tells* us something. And it is an anger that has a *right*. They *ought to* be angry when patriarchal arrangements leave boys and men alone, unloved, and emotionally deprived. It teaches them to channel their anger and invalidated emotions into violence. He knows by the end of the novel, patriarchy does *not* work in favor of them.

Gordie writes for a similar reason. Throughout *The Body*, he repeats, "The most important things are the hardest to say." He opens the book with that very line in a short piece that becomes the entirety of chapter one:

> The most important things are the hardest to say. They are the things you get ashamed of, because words diminish them—words shrink things that seemed limitless when they were in your head to no more than living size when they're brought out. But it's more than that, isn't it? The most important things lie too close to wherever your secret heart is buried, like landmarks to a treasure your enemies would love to steal away. And you may make revelations that cost you dearly only to have people look at you in a funny way, not understanding what you've said at all, or why you thought it was so important that you almost cried while you were saying it. That's the worst, I think. When the secret stays locked within not for want of a teller but for want of an understanding ear. (King 2018, p. 1)

Gordie becomes a writer as it is *the* outlet for him to express the masculinity that conflicts with patriarchal standards and expectations of his time. The "most important things" he writes about are his ability to desire love, that is things he is secretive of. Gordie regrets not telling Denny before he died that he loved him, because jealousy blinded that love. He also regrets not telling Chris he loved him before they parted ways after returning from their trip and before he was ultimately killed. He fears that "giving away" his secret, desiring love, will make him further ostracized from society. He would be seen as strange or (the term they use in the novel) a "pussy." To confess his feelings of the erotic, the "pleasure" to love and feel, is to admit he was indifferent to patriarchy. And if one admits they never cared for the laws imposed on them, they are reprimanded. Being an artist, a writer, is the most appropriate way to express *eros*—the only way Gordon believes he can without being judged or chastised. Thus, he completes his philosophical "becoming" of man-who-is-woman by finally allowing himself to feel the emotions he has buried since he was a boy which have been disregarded and ignored all this time.

Boyhood in these novels show us something novel. It tells boys and men that they do not need to abide by patriarchy to be loved and recognized. S.E. Hinton's *The Outsiders* and Stephen King's *The Body* are two novels that send that very message to its male readers. In fact, they *ought to question it.* And I believe writers and feminists such as hooks overlooked it. To truly wish for feminism, it is true that we must address not only women's anger and emotions, but men's as well. Patriarchy does not only hurt and victimize women, but men as well. Olufemi Taiwo has already identified and even proposed certain solutions to patriarchy. Philosophers such as Spelman, Lorde, and Cixous support further that these solutions begin in identifying and understanding the emotions, injustice, and indifference to law. In the case of addressing masculinity, *men's* emotions, injustice, and indifferences to the law. Which have been time and time again, discarded and made an afterthought in feminist discussion and literature. Men have shown it to us in small and subtle ways, and we have *still* discarded and overlooked it. The will to change is not only a demand for men, but for women as well. We need to have these conversations with men. We must keep encouraging this creativity, this novel form of masculinity, to be told and shared. To help them identify and validate their anger, for it is the anger of people who have been betrayed and wronged by a social framework that never cared for them in the first place, who have been told false truths about their identities. And we want to start this process by acknowledging they *already* have shown their will to change.[7]

End Notes

[1]"Red pill" is a term based off the film *The Matrix* (1999) where the protagonist takes the red pill to reveal the truth of his world. It is used by the "Manosphere" community to suggest that the "true" reality men face is misandrist and favoring women. The typical rhetoric of these types of media asserts that women must "return to being submissive" (UN Women.org 2025).

[2] Although I discuss the original novel versions in this article, their respective film adaptations' (*The Outsiders* by Francis Ford Coppola in 1983 and *The Body* under the title of *Stand by Me* by Rob Reiner in 1986) dialogues, interactions, and events of both stories (especially the specific details I mention in this article) are nearly identical. This paper's arguments can still be supported by their film counterparts. Interestingly, the very fact that these two novels had been given film adaptations (which have been widely received well in popular culture) enhances my argument that the "masculine media" bell hooks is critical towards can show a wide audience potentially non-patriarchal forms of masculinity.

[3] There is much evidence throughout the novel to show the Greasers expressing emotions but do not exactly know *how* to manage nor identify them appropriately. They do not seem to even know why they exactly do certain things. An instance of this is Ponyboy asking his friends and brothers why they all fight in gang rumbles and they do not give detailed reasons other than "for fun" or contempt for the Socs (Hinton 2006, pp. 100-106).

[4] Refer to Part I, 1.1; (Wistisen 2021, pp. 205-206)

[5] The other boys, Vern and Teddy, both pass in accidents. This ultimately makes Gordie's reflection of the events of the novel more somber and melancholic in tone. However, in the film adaptation, *Stand by Me*, while the film keeps Chris' death, Vern and Teddy do not die but simply grow out of the friend group. I could not find any explanation for this change, but regardless, the ending of Gordie being "alone" and the most successful out of all the boys still supports my argument.

[6] I put the term "women" in quotations here in scare quotes to remind readers once again *not* to take Cixous and mine's use of "feminine" and "woman" literally when discussing her philosophy. Cixous' notions on philosophical and artistic becoming can pertain to *anyone,* regardless if they are men and women.

[7] I would like to thank and acknowledge Dr. Katie Howard for directly supervising this project from outline to finale draft and for her guidance and teachings on both feminist philosophy and philosophy and the emotions—the foundations of this paper. I would also like to thank and acknowledge Dr. Jay Conway for his support throughout my pursuit of philosophy and introducing me to (and fostering my interest for) feminist philosophy, history of philosophy, and aesthetics—the heart that ties the foundations of this paper together. This paper would not be possible without them.

Bibliography

Aristotle. (2014) "On the Generation of Animals" in *The Complete Works of Aristotle: The Revised Oxford Translation, One-Volume Digital Edition,* pp. 1111-1218

Beals, Sandra. (2018) "Modeling Liberation: Audience, Ideology, and Critical Consciousness in S.E. Hinton's *The Outsiders*," *Children's Literature Association Quarterly* 43 (2) pp. 183-201

Cixous, Hélène. (1987) "Reaching the Point of Wheat, or a Portrait of the Artist as a Maturing Woman," *New literary history*, 19 (1), pp.1-21

Epictetus. (1904) "On the Things in Our Power, and Not in Our Power" in *Discourses*, pp. 1-5

Hinton, Susan Eloise. (2006) *The Outsiders* (Penguin Classics)

hooks, bell. (2004) *The Will to Change* (Washington Square Press)

King, Stephen. (2018) *The Body* (Scribner)

Lorde, Audre. (2012) *Sister Outsider: Essays and Speeches* (Berkeley: Crossing Press)

Spelman, Elizabeth. (1989) "Anger and Insubordination," in: *Women, Knowledge, and Reality: Explorations in Feminist Philosophy*, pp. 263-273 (Boston: Unwin Hyman)

Taiwo, Olufemi. (2020) "Stoicism (as Emotional Compression) is Emotional Labor," *Feminist Philosophy Quarterly* 6 (2)

UN Women. "What Is the Manosphere and Why Should We Care?" *UN Women*, 15 May 2025, www.unwomen.org/en/articles/explainer/what-is-the-manosphere-and-why-should-we-care.

Wistisen, Lydia. (2020) "Too Much Feeling: S.E. Hinton's *The Outsiders* (1967), Conflicting Emotions, Identity, and Socialization," *Children's Literature in Education* 52 (2), pp. 200-216

Animalism Does Not Entail Sex-Essentialism: A Defense of Trans Identity Following the Organism View of Identity

Daniel Garcia

Introduction

The anti-trans philosophical project is a crucial forerunner to fascist legislation which formalizes various forms of violence and abuse. The rhetoric of anti-trans bills relies on specific philosophical frameworks which presume the invalidity of trans identity as a methodological starting point. Consequently, anti-trans arguments often collapse upon further analysis of their logical structure. This is evidently the case in Moschella's 2021 article, "Trapped in the Wrong Body? Transgender Identity Claims, Body-Self Dualism, and the False Promise of Gender Reassignment Therapy". In this article, Moschella attempts to ward off trans identity claims through a flawed metaphysical argument which misappropriates animalism to advance a sex-essentialist account of personal identity. Using the "wrong body account" as an entry point, Moschella claims that trans identity relies upon Cartesian dualism to secure access to gender affirming medical care. Moschella proposes that a non-dualist metaphysical view, namely animalism, entails that trans identity claims are indicative of a fundamentally disordered perception of reality. Moschella then advances a moral argument against gender affirming care, asserting that the proper treatment for distress due to false belief must be psychological treatment and not body alteration.

In this paper, I will debunk Moschella's metaphysical claim that animalism entails sex-essentialism. I argue that animalism supports trans identity claims such that reproductive sex is not essential to human identity. Consequently, Moschella's metaphysical argument does not support her moral claims and trans identity claims remain unassailable. In section 1, I explain Moschella's argument and reconstruct the logical structure of her work. In section 2, I discredit the metaphysical foundations of Moschella's claims by granting animalism and demonstrating how sex-essentialism does not follow. In section 3, I discard Moschella's flawed moral argument.

Section I: Moschella's argument

The primary thesis of Moschella's article is that transgender identity claims entail that a person is not numerically identical to their own body, a view that Moschella regards as metaphysically impossible. Moschella begins with an analysis of the wrong-body account (WBA), a narrative framework often used to express a supposed incongruence between one's gender identity and one's physical body, resulting in feelings of being "trapped in the wrong body". Moschella takes this to be a literal metaphysical claim which espouses mind-body dualism, the view that a person's mind is separate and distinct from their physical body. Moschella demonstrates the errors of mind-body dualism and consequently discredits the ethical grounds for gender-affirming care supported by this view.

Following contemporary philosopher Eric Olson, Moschella argues for animalism, a biological view of personal identity which posits that the human person (i.e. the self) is identical with the human animal (i.e. the body). Animalism rejects the Lockean view of personal identity, which theorizes that the human person (i.e. the self) is identical with their consciousness (i.e. the mind). Moschella, having rejected this view, argues for an even narrower biological view of personal identity, whereby the self is essentially the sex (i.e. the body's overall organization to reproduce as male or female). This view is consistent with the contemporary gender-critical view, which rejects the concept of gender identity altogether and purports a sex-essentialist view of personal identity (i.e. the self is the sex, and nothing else). By rejecting the psychological view of personal identity, Moschella defuses any gender identity claims which suggest an incongruence between the body (i.e. sex) and the mind (i.e. gender). With this, Moschella believes to have conclusively discredited trans identity claims.

Following her metaphysical thesis, Moschella concludes with a troubling picture of trans identity and gender dysphoria. Moschella's account frames trans people as psychologically disordered, such that trans identity claims must involve a fundamental misperception of reality. Likening gender dysphoria to anorexia, Moschella argues that bodily alteration is not the proper medical treatment for distress which appears due to false belief. Moschella maintains the gender-critical view that gender affirming care is premised on an illusion (that one may change the sex, i.e., the body), rather than a metaphysically sound understanding of reality (i.e., the self is essentially the sex, and nothing else). Moschella corroborates this with detransitioner testimony and statistics which correlate gender dysphoria with other mental health issues, arguing that gender affirming care fails to address the underlying causes of distress. Thus, Moschella arrives at her ethical thesis, that trans identity warrants psychological treatment and not bodily alteration. Since trans identity is a false belief which causes distress, Moschella argues, the ethical treatment must address the root cause of suffering, which is psychological, rather than promote false belief with gender affirming medical care.

The logical structure of Moschella's argument is represented with the following six points:

1) Trans identity claims rely on the wrong-body account (WBA), a metaphysical claim which presupposes Cartesian dualism.
2) Animalism is true (i.e. we are numerically identical to our bodies) and dualism is false.
3) Therefore, the self is the sex, and nothing else.
4) But, trans people experience body-self mismatch.
5) This must be a false belief due to a misperception of reality (i.e. psychological error).
6) Trans identity warrants psychological treatment and not gender affirming care (i.e. bodily alteration).

Claims 1-3 contain Moschella's metaphysical thesis, while claims 4-6 rehash the conventional anti-trans view of trans identity. Arranging her argument in this structure elucidates how the latter is entailed by the former; or rather, how the latter necessitates the former. In other words, it is clear from Moschella's poor philosophical methodology that her work is primarily concerned with securing a convincing anti-trans ethical position, which she achieves by advancing a contentious metaphysical argument. By analyzing each of Moschella's claims in the following sections, I will demonstrate how Moschella fails to deliver a logically consistent, structurally valid, and intellectually sound philosophical argument for the anti-trans ethical view. Notably, I will grant claims 2 and 4 as they both seem to be true and provide a common ground from which I may present a defensive pro-trans argument which is non-dualist.

SECTION II: MOSCHELLA'S FLAWED METAPHYSICS

Primarily, we may throw out premise 1 and thus dismiss Moschella's entire argument. It is entirely obvious that Moschella's argument rests on the disingenuous assumption that the WBA must be a literal metaphysical claim entailing dualism. This assumption is easily disputed for two reasons. One, not all trans people subscribe to the wrong-body account. Moschella herself acknowledges that the WBA is controversial, especially in the field of trans philosophy, for its problematic implications which simultaneously leave out members of the trans community while assimilating all trans narratives into one specific paradigm.[1] Two, the WBA is not a literal metaphysical claim.

It is a rhetorical device. Moschella provides no evidence that the WBA must be taken as a literal metaphysical claim and indeed commits a category error by classifying it as such. The WBA may instead be taken as a recurring motif in trans identity narratives, which speaks to how trans communities have come to articulate complex lived experiences rather than suggest a belief in Cartesian dualism. Moschella's argument against the rhetorical WBA does not entail a takedown of every trans identity claim. Her entire argument, being premised on a false claim, falls apart upon rejection of premise 1.

Next, we may reject premise 3 because it does not follow premise 2. Animalism does not entail sex-essentialism and Moschella fails to justify the claim that the self is essentially the sex and nothing else. Moschella presents a distorted, revisionist version of animalism and personal identity to achieve her ends.

Animalism is a metaphysical view which posits that human people are essentially animals, such that one's existence is coextensive and inseparable from the human biological organism. The emergence of animalism as a concept may be traced through the early modern philosophical discourse sparked by René Descartes's *Meditations on First Philosophy.* Widely regarded as the father of modern Western philosophy, Descartes's *Meditations* establish mind-body dualism through an interrogation of the nature of self and reality. As a substance dualist, Descartes argues that bodies (i.e. material objects which take up physical space) are distinct from minds, or "thinking things" (i.e. abstract, purely intellectual beings). Consequently, the self must be identical with the mind and not the body (Descartes 1993, pp. 19-20).

Descartes's substance dualism puts forward two distinct identity conditions. For bodies, continuity of existence is guaranteed by the underlying substance which contains the object's physical properties. This concept of identity is demonstrated through the example of beeswax presented in the second meditation (Ibid., pp. 21-24). In its initial state, a piece of beeswax contains certain perceivable properties such as its smell, shape, color, texture, etc. When placed near a fire, those properties change. The wax melts, loses its color and shape, becomes liquid and hot. We no longer perceive the same properties but still regard the wax as the same piece of wax. Descartes argues that the piece of wax retains its identity, not because of its perceivable qualities, but due to some underlying substance which carries the object's properties. By contrast, the existence of the self is only guaranteed by thinking. Since the self is a purely intellectual thing, one may only be certain that the self exists when one understands that one exists. The mind-body problem emerges as a consequence of this dualism (Ibid., pp. 19). Descartes provides no conclusive solution for how the mind, which is purely intellectual, can interact with the body, which is purely physical. There is obviously some causal interaction between these two radically different substances, and it seems intuitively true that we are identical to our bodies (i.e. dualism is false and animalism is true).

Locke, a British empiricist, engages with the concept of identity in *An Essay Concerning Human Understanding*, beginning with his theory of knowledge. Rejecting Descartes's doctrine of empirical skepticism, Locke argues that all ideas, both simple and complex, come from sensational experiences (Locke 1961, pp. 17-36). When we reflect upon our rich and complex sensational experiences, we utilize type-identical language to name particular objects as we perceive them. This process of reflection and abstraction misleads us to believe in the existence of abstract concepts which, in our minds, correspond to the object itself. Consequently, Locke rejects the existence of substances, denying that some underlying substrate exists which guarantees an object's identity.

Instead, Locke argues that an object is merely a pile of atoms. The object's identity—what makes the object self-identical over time—can only be guaranteed by its mass. If you add or remove a single atom from the object, that object is no longer self-identical. This is the principle of individuation (Ibid., 64). From this view, Locke would argue that Descartes's piece of wax is an individual object insofar as it is one distinct mass. The wax's particular properties (such as its shape, smell, texture, etc.) are perceivable qualities which exist due to the arrangement of atoms which comprise the material object (Ibid., pp. 52-60). Were Descartes to break off a piece of the beeswax, the two pieces of mass would no longer be identical to the original piece of wax.

Consequently, Locke's materialism unsettles the concept of organic identity. Living creatures differ from other material objects such that an organism's individual mass is constantly changing. Locke addresses this with the principle of organic identity (i.e. what makes one organism distinct and self-identical over time) (Ibid., 65). Essential to each living organism is the body's overall organization of a system for sustaining life. According to Locke, adding and removing mass is inconsequential to the organism's individual identity, so long as the system for life remains continuous. Thus, Locke establishes different identity conditions for masses of matter and for living bodies.

The individuation of organic identity guarantees the existence of the human body as one continuous system of life, but fails to secure the existence of the human person as a rational being distinct from other living creatures. Thus, the problem of personal identity emerges. Locke, in fear of reducing humans to the same existential status as animals, introduces the concept of consciousness. Locke argues that personal identity, unlike brute or organic identity, is determined by consciousness. Following Descartes's assertion that the self is essentially the thinking thing, Locke argues that, because thinking is always done by virtue of consciousness, the self is essentially the consciousness (Ibid., 67). However, consciousness alone cannot guarantee identity—sameness over time—because one's consciousness is always changing. According to Locke, the continuity of consciousness is guaranteed by memory. Thus, Locke establishes that a person's existence is co-extensive with their consciousness and guaranteed by memory (i.e. psychological continuity). Consequently, this view implies that some humans are brute animals but not people insofar as memory is fallible and consciousness is not always continuous. For instance, within Locke's view, the fetal body which develops prior to one's conscious memory is not numerically identical to the adult human person. In other words, if you don't remember being a fetus (i.e. there is no psychological continuity), then that fetus is not you. Therefore, Locke cannot guarantee personal identity, i.e. sameness over time, with consciousness alone.

Animalism intervenes upon the problem of personal identity by rejecting the psychological view and embracing the organism view. The main proponent of this view in modern philosophy is contemporary philosopher Eric T. Olson, who writes *An Argument for Animalism*. Animalism applies Locke's principles of organic identity to guarantee the existence of the human animal through biological continuity. In other words, Olson argues that humans are animals insofar as our human existence is obviously co-extensive with our organic existence (Olson 2003, pp. 8-12).

Importantly, animalism is not the view that we are identical with our bodies. Olson avoids claims which identify the animal as the body because the truth of such claims depend entirely on what it means for something to be someone's body (Ibid., 7). Olson illustrates this by supposing that someone might have a partly or wholly inorganic body, writing, "If you cut off an animal's limb and replace it with an inorganic prosthesis, the animal just gets smaller and has something inorganic attached to it" (Ibid., 6). In keeping with Locke's view of organic identity, adding or removing mass is inconsequential to the identity of the organism as a system of life. Therefore, in the example above, the animal remains self-identical although the body has changed or contains inorganic parts. This example clarifies that animalism is solely concerned with delineating the persistence conditions of the human animal as a living organism. In other words, animalism does not say that the self is essentially the body.

Additionally, animalism does not say that we are merely animals. Being an animal does not rule out other properties which do not follow from being an animal. For instance, Olson argues that Descartes was a philosopher, but not merely a philosopher. He was also a mathematician, a Frenchman, a Roman Catholic, etc. These properties are not ruled out by Descartes being a philosopher. Olson clarifies, "Animalism does not imply that we have a fixed, "animal" nature, or that we have only biological or naturalistic properties, or that we are no different, in any way, from other animals" (Ibid., pp. 7-8). Olson argues that although there is indeed a vast difference between human animals and other living creatures, especially in our capacity for rational thought and moral practice, humans are nevertheless animals.

Animalism answers the problem of personal identity by clarifying the persistence conditions for our existence as human animals. As human animals, our existence is identical with the existence of the human organism as an organized system of life. Animalism rejects the Cartesian notion that thinking is co-extensive with existence, which Locke adheres to with the concept of consciousness and memory. These claims do not provide for other fundamental properties of personal identity, nor does animalism exhaustively define what constitutes a person. Put simply, although someone in a vegetative state might cease to exist as a conscious person, one would not ordinarily suppose that the living organism is numerically unidentical to the human that existed prior (Ibid., 22). This is how animalism demonstrates that biological continuity, not mental continuity, is a necessary condition to guarantee the numerical identity of a living organism.

Moschella's reconstruction of animalism is therefore distorted and intellectually dishonest. The first breakpoint in Moschella's animalist argument appears immediately after she explains animalism, having referenced the same 2003 paper by Eric Olson, to state: "What is important for the purposes of this article is simply that, regardless of whether or not there are aspects of our identity that go beyond our biological identity, our biological identity is essential and intrinsic to our personal identity. In other words, I *am* my body, even if there are aspects of me that transcend the bodily dimension of my being." (Moschella 2021, pp. 787-788)

Rejecting dualism and embracing the animalist view does not entail that the self is necessarily the body. Animalism, as explained by Olson and supported by other contemporary philosophers, explicitly avoids identifying the human animal as the body. One simply cannot justify the claim that the self is the body using animalism, even for the purposes of Moschella's article. Furthermore, according to animalism, our biological identity—i.e. our organismal continuity as human animals—is explicitly *not* essential to our personal identity, such that our biological identity may be continuous even when our existence as ordinary people might be suspended. Animalism specifically serves to unsettle personhood as the condition which guarantees our continuity as living organisms. Therefore, Moschella's claim that the self is essentially the body remains unjustified, rendering her argument logically invalid.

Moschella then goes on to further misappropriate animalism to advance a sex essentialist view of biological identity, writing, "For if I am my body, then the characteristics of my body, including the complex and pervasive ways in which my body is organized to play a female role in reproduction, are characteristics of *me*. In other words, if my body is female, *I* am female." (Ibid., pp. 788) Here, Moschella clearly misappropriates the language of animalism to present a distorted view of organic identity, revising the terms to shoehorn reproductive sex as a necessary condition of life. Animalism does not entail that one bodily characteristic, i.e. the body's overall organization for reproduction as male or female, is essential to existence as a human animal. The body's sex is in fact contingent, such that the identity conditions for the human animal solely concern one's persistence through time as a single organism. Though being male or female is compatible with being a human animal, the male/female sex identity is in no way essential to organic identity. Were sex to be essential to human animal identity, the persistence conditions for males and females would be necessarily different. However, as it were, the persistence conditions are no different for any living organism, such that the same principle of organic individuation may be applied indiscriminately to males, females, humans, horses, and even single-celled organisms. Therefore, neither reproductive sex nor species membership are necessary or essential to one's organic identity as a living animal.

Evidently, premise 3 serves as a crisis point for Moschella's argument. Moschella begins with the ethical conclusion that trans identity and gender affirming care are wrong and thus stipulates a metaphysical thesis which would satisfy this conclusion. Moschella's attempt to justify sex-essentialism using animalism demonstrates the academic bankruptcy of her work. This move is not unique to Moschella, as this form of question-begging has become standard practice in anti-trans circles. For instance, Moschella references a 2008 book by Patrick Lee and Robert P. George entitled *Body–Self Dualism in Contemporary Ethics and Politics*. This work similarly attempts to misconstrue animalism in a critique of body-self dualism as an entry point to espouse Catholic and evangelical moral principles regarding issues such as drug-taking and sex outside of marriage. A reviewer writes of Lee and George's book, "This work is clearly a case of the authors beginning with their ethical conclusions, and then casting about for whatever metaphysics will corroborate their case" (Robinson 2009, pp.155-157).

Section III: Moschella's Account of Trans Identity

Having discredited the metaphysical foundations of Moschella's argument, we may consequently discard her ethical conclusions regarding trans identity beginning with a re-examination of premises 4-6.

First, we may grant premise 4 because it seems true. Trans people do exist, accompanied by some phenomenological experience which seems to motivate transition and is often articulated through a framework of body-self mismatch.

Premise 5 fails in that she attempts to present a monolithic account of this phenomenon, assimilating a diverse set of experiences, testimonies, statistics, and philosophies into a single narrative which paints trans people as psychologically disordered. Regardless of Moschella's metaphysical view, philosophical disagreement about the nature of trans phenomena does not entail psychological error nor false belief. In other words, premise 5 is simply not true nor defensible because an alternative metaphysical view of personal identity does not grant Moschella the authority to deem trans people as fundamentally disordered.

Premise 6 reveals the true agenda for Moschella's article: to present a metaphysical justification for a specific moral thesis. With premise 6, Moschella asserts what we ought to do about trans phenomena, namely, that trans people ought to be denied gender affirming care and instead convinced to accept a sex-essentialist view of personal identity. This argument presents a distinctly Catholic view of medical ethics, such that doctors should enforce a specific worldview by regulating access to various medical technologies.

We may discard Moschella's ethical view entirely as it does not follow from her metaphysical argument.

Section IV: Conclusion

Moschella's article fails to present a convincing metaphysical argument against trans identity claims and gender affirming care. Premises 1-6 are fraught with academic dishonesty, structural errors, and logical fallacies. Premises 1-3 rely on a misleading metaphysical thesis of sex-essentialism which does not follow from Moschella's premise of non-dualist animalism. Premises 4-6 present a moral argument, in line with the contemporary gender-critical view, which attempts to paint a disparaging view of trans people according to a flawed metaphysical thesis. Though Moschella attempts to provide a philosophical basis in this paper, anti-trans arguments remain metaphysically unjustified.

Bibliography

Moschella, Melissa. (2021). “Trapped in the Wrong Body? Transgender Identity Claims, Body-Self Dualism, and the False Promise of Gender Reassignment Therapy,” *The Journal of medicine and philosophy*, 46(6), pp. 782–804.

Descartes, René. (1993) *Meditations on First Philosophy.* Translated by Donald A. Cress. (Indianapolis: Hacklett Publishing Company)

Locke, John. (1961) “An Essay Concerning Human Understanding (Abridged),” in: R. Taylor (Eds), *The Empiricists*, pp.7-133 (New York: Anchor Books)

Olson, Eric T. (2003) “An Argument for Animalism,” in: R. Martin & Barresi (Eds), *Personal Identity*, pp. 318-334 (Oxford: Blackwell readings in philosophy)

Robinson, Brian. (2009) “Body–Self Dualism in Contemporary Ethics and Politics,” in: *Journal of Church and State,* 51(1), Winter, 2009, pp.155–157

On The Ubiquity Of Violent Media And The Nature Of Its Influence On Moral Character

Giovanni Galvan

Introduction

Violence has been depicted prominently throughout human history. From cave paintings to cinematic recreations, violent media has persisted ubiquitously as a means of human expression that continues to grab our attention today—but why? It seems rather counterintuitive that media depicting fictive acts of violence that—whether historical or entirely fictional—would otherwise be harmful (or possibly fatal) if enacted in reality but can conversely be considered a captivating form of entertainment—but never mind that; if we treat such media as entertaining (as opposed to distressing), are there still aspects of its nature that remain harmful? If so, how and why are these aspects harmful?

To preface, the type of medium through which the violent depictions discussed in this paper are expressed will be solely visual. The following arguments, texts, and opinions were conceptualized within the context of such media in mind. Violence is defined as "intentional physical harm", and visual media refers primarily, but not limited to, film and television. It is acknowledged that the attraction to violence being discussed is not a universal experience but is nonetheless presupposed to be a prevalent sentiment held by audiences of violent media. 'Violent media' will not refer to all media that feature violence (e.g., *Food, Inc.*)[1], but media that feature depictions of violence and are themselves considered problematic as a result of these depictions (e.g., *Texas Chainsaw Massacre*)[2]. The difference that separates the latter from the former lies in its overall characterization as violent, rather than just being treated as media featuring depictions of violence.

In Section I, I provide a brief overview of the pertinent discourse on the (alleged) effects of fictive depictions of violence in recent years. I will then follow with an introduction of philosophers, Paul Duncum and Ian Stoner, as well as the opposing arguments articulated in their respective articles. In Section II, I will give a more comprehensive and concise outline of Duncum's argument. In Section III, I will—similarly—present a more comprehensive yet concise outline of Stoner's argument. In Section IV, I articulate my position on the issue while referencing relevant aspects of each author's argument, whether in support or opposition. Finally, I conclude the article with a brief overview of the most important points of each section.

SECTION I: CONTEMPORARY ATTITUDES TOWARD VIOLENT MEDIA

Discourse regarding violent media and its implications on moral character has focused on its supposed effects on moral psychology. Such discussions have held an important place in the film and television industries since their inception. Contemporary discourse has focused on various forms of the same argument: the ARA (Argument from Reactive Attitudes), which—in the context of violent media—posits that consuming media harmful to our moral psychology is immoral; violent media is harmful to our moral psychology; and, therefore, immoral to consume. I categorize the views behind these discussions into the following two categories: pro-ARA and anti-ARA. Views categorized under pro-ARA support the idea that, according to the ARA, exposure to depictions of violence in media is harmful. Views categorized under anti-ARA oppose the idea that the ARA is unsound and that exposure to depictions of violence in media is harmful.

The first article I discuss is Paul Duncum's "Attractions to Violence". Duncum's article is intended to represent the pro-ARA side of the discussion. Although Duncum does not directly refer to the ARA, the article centers around another form of the same argument: Violent media induces fear and anxiety, which leads to a further corruption of our moral psychology. These three consequences engender the harm that violent media is allegedly responsible for. Duncum argues the following: there is a current increase in violence and its exacerbating factors throughout contemporary visual media; violent media gives rise to and maintains a culture of fear and anxiety; and there is a cyclical and symbiotic relationship between violent media and the culture of fear and anxiety that perpetuates this harm.

The second article I discuss is Ian Stoner's "A Defence of Gorey Fictions". Stoner's article is intended to represent the anti-ARA side of the discussion. Stoner directly refers to the ARA throughout his article, and it is the main subject of his argument. Stoner argues that the ARA is unsound because of flaws in its second premise. Specifically, the assumption that fictive depictions of violence encourage the corruption of our reactive attitudes is flawed because it is inconsistently applicable; the same acts of mediated violence can be featured in a film that the ARA considers morally questionable and a film that is considered ARA-compliant, while only the former receives social repercussions.

Section II: Why Violent Media Is Harmful

Paul Duncum's "Attractions to Violence" explores audiences' seemingly widespread tendencies toward violent media via a nuanced conceptualization of these depictions. He lays out his argument against the ARA by first acknowledging a few points, describing four kinds of violence in media, individually assessing their moral and social implications, and concluding with his perspective on the state of violent media in American culture and how we should respond educationally or socially.

2.1: Preliminary Considerations

Duncum begins his article by acknowledging the following: attraction to violence isn't a universal experience; the most popular form of entertainment media (e.g., film and television) isn't violent; despite the first two acknowledgements, violence persists as a prominent element in all forms of media. He establishes these points to describe the context in which his argument was made. Furthermore, Duncum presupposes that the effects of violent media are of significant concern for parents and educators, and that violence is defined as "intentional physical harm".

Duncum proceeds to acknowledge the uncertainty surrounding the nature of our attraction to violence. However, he attributes its attractiveness to appeals to safety and mood management. The appeal from safety ascribes the attractiveness of violence in media to its utility as a means to an end. For most violent media, violence serves the purpose of being a literary device—an enhancer of the media's dramatic points rather than being a point in itself. In addition to this, violent media provides a safe vista from which audiences may experience the dangerous, adrenaline-inducing situations that our civilized society has protected us from. The appeal to mood management immediately follows from this, proposing that such a privilege allows audiences to experience the emotions characteristic of such thrilling situations.

Duncum moves forward by establishing the goals of his paper as answering the following three questions: (1) What is it about displays of violence that we find appealing? (2) Why is the intensity of violence in violent media increasing? (3) What are we doing to ourselves when we are exposed to such violent media? He considers the answers to these questions imperative to the understanding of the harm brought about by violent media, but recognizes the complex nature of violence in media. Duncum addresses this concern by reevaluating mediated violence across different categories based on a depiction's service as a means.

2.2: FOUNDATION

Duncum presents his nuanced account of the nature of violence in visual media, categorized under the following 4 categories: (1) Comic; (2) Transgressive; (3) Retaliatory; (4) Gratuitous (Duncum 2006, pp. 21-24). Comic violence refers to any violent—although seemingly foolish—act whose purpose is to satirize serious depictions of violence. Transgressive violence refers to any violent act performed with the purpose of breaking social, moral, or legal rules—commonly utilized as a medium through which audiences can judge a character's morality. Retaliatory violence refers to any violent act executed in retaliation for a previous offense. Gratuitous violence refers to any violent act undertaken for its own sake (i.e., as an end in itself). Duncum believes that each type of violence in the media warrants an individual assessment of its purpose and social implications, with the aim of narrowing our concern. He measures these constructs according to their respective need for social or educational intervention.

Comic violence differs from the three other kinds of violence in that its satirical nature serves a sobering end; it is unique in its purpose and goal. Duncum cites the following media as examples of such violence: cartoon violence (e.g., *Tom and Jerry*), slapstick violence (e.g., *The Three Stooges*), and theatrical violence (e.g., *WWE*). To aid its understanding, Duncum compares the role of comic violence to that of violence featured in medieval carnivals—violence whose purpose is to lighten the "vicissitudes" of life. Ultimately, given its positive intentions and cathartic effects, social/educational intervention is not considered necessary in cases of comic violence (Ibid., pp. 24-25).

Transgressive violence is a common kind of violence in visual media, especially in genres that involve action or redemption. As stated previously, it refers to unlawful or immoral violence that acts as a medium for a fictional agent's moral judgment. Duncum considers transgressive violence to be a common theme among violent video games, placing protagonists in a moral dilemma of utilizing unpleasant means to achieve a higher end. He attributes its emotional appeal to the audience's exercise of moral judgement and to its counterintuitive nature. Audiences enjoy the outside perspective from which they can engage in the preemptive judgment of a fictional agent and respond positively and accordingly to its further corroboration—an experience referred to as *forepleasure* (i.e., pleasure experienced in anticipation of a protagonist's retribution for their subjectification to transgressive violence). Duncum considers social and educational interventions potentially problematic due to the "boundary-testing" nature of adolescence (Ibid., pp. 25-27).

Retaliatory violence is an attractive form of mediated violence due to its emotionally arousing nature. Duncum considers retaliatory violence a common theme in shooter games, given their narratives' reliance on it. Although it may appear similar to transgressive violence, retaliatory violence refers to violence committed by a fictional agent in *response* to an offense taken against them; essentially, violence in response to transgressive violence. Furthermore, its effectiveness is heightened by the following psychological principle: one's level of arousal prior to an emotional experience determines the intensity of that experience. Duncum considers that the effectiveness of social/educational interventions is limited by the appeal of retaliatory violence to a fundamental principle of human psychology: that retaliation is cathartic (Ibid., p. 27)

Gratuitous violence is the senseless form of mediated violence due to its lack of purpose other than being an end in itself. Duncum cites the following media as hosts of such violence: slaughterfest films (e.g., *Final Destination* franchise), splatter films (e.g., *Texas Chainsaw Massacre*), and shoot-em-up games. Gratuitous violence is unique in that it is a kind of violence disproportionately favored by adolescent males. Duncum cites research suggesting that such depictions of violence provide this demographic with the opportunity to overcome the pressures and fears placed upon them by gendered roles (Sparks & Sparks 2002, pp. 114-126).[3] Furthermore, Duncum claims that the most attracted members of the demographic are isolated, later attributing this relationship to their limited access to high-sensation activities. He concludes by emphasizing that social/educational interventions are critical and essential but are unlikely to be effective due to our media culture's increasing featuring of gratuitous violence (Ibid., p. 28)

2.3: ARGUMENT

Duncum considers gratuitous violence—as the only form of mediated violence whose effects have the potential of receiving effective social/educational intervention—to be harmful for audiences to consume in excess, especially for individuals with antisocial tendencies. However, despite this harm, he argues that banning media containing gratuitous violence would be practically ineffective in preventing violence in the real world. Subsequently, he proposes that educating audiences in how to manage the fear and anxiety instilled by such media is the best practical and direct response to dealing with mediated violence.

To provide support for his claims of mediated violence increasing in both prevalence and intensity, Duncum first analyzes the history of violence in the media and investigates the factors that have exacerbated its growth. Violence in media was first regulated by the Motion Picture Association of America Code in 1930, but has become increasingly violent since at least the 1960s. Ever since then, violence has maintained a continuous presence in the media, but its intensity has increased dramatically—but why? Duncum believes that it is due to a combination of chronic habituation and mediated violence's efficacy as a subconscious attention-grabber (Ibid., p. 29).

Habituation raises audiences' threshold for violence and lowers the intensity of their reactions. Duncum cites research supporting that a negative relationship between exposure to violence and the intensity of future reactions exists (Cantor 2002, pp. 197-219; Horsley 2002, p. 43; Zillman 1988, p. 182).[4] This effect not only affects the individual but also the cultures it plagues. As a result, the media's widespread dissemination of mediated violence contributes to a mass habituation with violence. Similarly, the captivating nature of mediated violence presents itself to media producers as a subliminal and reliable means of capturing audiences' attention, thus providing a strong incentive for its use. Furthermore, the widespread use of violence in media induces mass habituation, which, in turn, leaves the minds of audiences demanding intensifying violence to continue captivating their attention; thus, a self-sustaining, predatory cycle harmful to our moral psychology is born. In a civilized society built by rational agents, the defiant, thrilling nature of the irrational can be found appealing (Ibid., pp. 31-34).

Duncum concludes his article with his final argument regarding how one should respond to mediated (especially gratuitous) violence and its effects on the individual, and how we should respond collectively. He begins his final argument by establishing that violent media induces fear and anxiety within the individual, and—en masse—the cultures in which they are consumed. In response, Duncum advises that, as an individual, we should focus our efforts on dealing with the real issue at hand: the fears and anxieties engendered by mediated violence. He lists three steps on how to deal with these fears and anxieties: (1) Focus efforts on the symbiotic aspect of the cycle between them and mediated violence; (2) Utilize rationality and evaluate the nature and validity of one's fears and anxieties; (3) Treat valid and rational fears and anxieties with genuine concern and respond accordingly. As a collective, Duncum suggests that we provide educational interventions for those unaware of the relationship between mediated violence, marketing, and our moral psychology (Ibid., pp. 34-35).

SECTION III: WHY VIOLENT MEDIA ISN'T HARMFUL

Ian Stoner's "A Defence of Gorey Fictions" explores the common assumption that consuming violent media is harmful and morally wrong. Stoner introduces the Argument from Reactive Attitudes (ARA) as an argument from which those who definitively hold such an assumption posit that we have a duty to maintain healthy moral and psychological responses, and that consuming violent media violates this duty by corrupting said responses. Thus, films whose depictions of violence are suspected of harming our moral psychology are also considered immoral to consume intentionally, and films whose depictions of violence are considered safe for our moral psychology are also considered permissible to consume intentionally. Additionally, the ARA is resistant to empirical critique because of its second premise, which offers a vague definition of corruption that allows even the slightest empirical change (empirically unmeasurable) in one's reactive attitudes to support it (Stoner 2002, pp. 511-527).

Prior to his main argument for why the ARA is false, Stoner explains that its second premise is questionable at best and requires an accepted, relevant distinction between the depictions of violence in films considered to be immoral by the ARA and films where the same depictions would not be suspected of being immoral (Ibid., p. 514). After establishing that the second premise is questionable, Stoner states that the ARA is to be considered unsound as long as there is no known relevant difference between the mediated violence featured in both kinds of films; thus, relieving violent media of any moral concern.

3.1: PRELIMINARY CONSIDERATIONS

For the sake of clarity, Stoner's account of the ARA and its premises is outlined below:

(P1) It is morally wrong to view features of a work of art
that encourage the corruption of reactive attitudes necessary
for the development of a well-functioning moral psychology

(P2) Taking pleasure in horror violence encourages the corruption of our reactive attitudes

∴ Therefore, it is morally wrong to intentionally view works of art depicting horror violence

Stoner provides clarification on which media are being discussed, presenting definitions and limitations of what the media may contain. For ease, I will use the term "questionable media" to refer exclusively to media whose depictions of violence are considered morally questionable and, thus, harmful (e.g., *Texas Chainsaw Massacre*, *Blood Feast*); and, "permissible media" to refer to media whose depictions of violence are morally permissible by the ARA and, thus, safe (e.g., *A Clockwork Orange*, *Pulp Fiction*); and "gore" to refer the depictions of violence present in both kinds of media and of similar severity. Stoner describes the current consensus on questionable media as morally concerned, but allowed to remain due to its censorship being of moral concern as well. Furthermore, he considers films that warrant defense from the ARA as "gore films with fictive representations of violence, suffering, and death". His use of "gore films" in this statement is referentially equivalent to my use of the term "questionable media". Therefore, the depictions of violence in media discussed in his article must be fictive recreations of violence, suffering, and death; thus, films containing depictions of real violence, suffering, and death are not the subject of discussion.

3.2: Foundation

Stoner clarifies that, regardless of the appeal that the ARA may have to those concerned with gore's impact on our moral psychology, the argument's second premise remains susceptible to being false. To cement his skepticism with the ARA's second premise, Stoner considers three candidate differences between the gratuitous violence featured in questionable media and permissible media: (1) artistic merit; (2) meaningfulness; and (3) audience orientation. These differences are attempts at answering the question, "If the fictive depictions of suffering and death featured in [questionable media] are considered dangerous to our moral psychology, why is it that the presence of these same fictive depictions within other mediums of art is left exempt?" Thus, Stoner intends to maintain the ARA's susceptibility by explaining why these candidate differences would and do not suffice (Ibid., pp. 514-515).

The answer from differences in artistic merit posits that questionable media and permissible media differ in their ends. Essentially, questionable media lacks the goal of exemplifying artistic prowess that permissible media aims towards. Using *Titus Andronicus* by Shakespeare as an example, Stoner explains that neither an excellent nor a poor production of the play would alter the play's violent content's impact on the viewer's moral psychology. If an excellent production of *Titus Andronicus*—setting the perfect example of artistic achievement by recreating the intensity and passion characteristic of its violent content—is considered permissible under the ARA, then a lesser quality production of the same play—featuring a less intense and passionate recreation of its violent content, exemplifying poor artistic achievement—would also be expected to be found permissible under the ARA. Thus, it follows that artistic merit is not a distinguishing difference between questionable and permissible media; if it were, a lack of artistic merit—as was the case in Stoner's example of the poor production of *Titus Andronicus*—would have deemed it morally questionable (Ibid., pp. 515-516).

The answer from differences in meaningfulness suggests that questionable and permissible media differ in the underlying nature of their gory content. The gore featured in questionable media lacks meaning and is only an end in itself; in contrast, permissible media only feature gore that serves as a means through which meaning can manifest. Stoner explains that most questionable media—especially the most worrying—feature narratives that tie the violent content of the films to further goals explored throughout. Using *Schindler's List*, Stoner explains that a showing of the first half of the movie—leaving only the gore and none of the deeper meaning presented in the latter half—would not revoke its permissibility. (Stoner 2020, p. 517) For the minority of questionable media that lack such meaning, this absence would not suddenly bring harm to viewers' moral psychology. Thus, it follows that meaningfulness is not a distinguishing difference between questionable and permissible media; if it were, a lack of meaningfulness—as was the case with Stoner's example of the cut *Schindler's List*—would have deemed it morally questionable (Ibid., pp. 517-518).

The answer from differences in audience orientation claims that questionable and permissible media differ in the intentions of their emotional appeal. Simply, the former differs from the latter in that their intention is to encourage a form of *vicarious sadism*—the finding of pleasure in the gore through the eyes of the perpetrator—as opposed to *paradoxical pleasure*—the finding of pleasure from the pain experienced as a result of witnessing the gore. Using *Paranormal*

Activity as an example, Stoner directs our attention to the use of fear and disgust in its marketing, where the movie's appeal lies in the converse promise of these emotions. The appeal of most questionable media lies in the promise of fear and disgust. Furthermore, fear and disgust are emotions that would not elicit in someone who was experiencing vicarious sadism (as a sadist would instead experience pleasure directly from the sadistic content), but would in those who are experiencing paradoxical pleasure (as such an experience is derived from the fear and disgust initially elicited upon witnessing the sadistic content). Thus, it follows that audience orientation is not a distinguishing difference between questionable and permissible media; if it were, then vicarious sadism would be a much more prevalent appeal for audiences of the former, rather than paradoxical pleasure maintaining dominion over audiences of both the former and the latter, ultimately acting as a common ground between them (as opposed to acting as a distinguishing difference) (Ibid., pp. 518-520).

3.3: ARGUMENT

As Stoner had previously stated, there must exist a unique difference between the depictions of suffering and death (i.e., gore) in questionable and permissible media for ARA's second premise to hold merit. However, as he has evidently established, there is no accepted, unique, and relevant difference between the gore in questionable and permissible media. Thus, Stoner considers this lack of an accepted difference sufficient for declaring a *provisional* conclusion for his argument: that there is no merit to the application of the ARA on the supposed immorality of the gore and mediated violence of questionable media (Ibid., p. 520)

SECTION IV: FINAL THOUGHTS

As my overview of these articles comes to an end, I would like to commend their content and authors for their comprehensive and intricate arguments. Duncum's "Attractions to Violence" and Stoner's "A Defence of Gorey Fictions" contributed significantly to the contemporary discourse surrounding violent media, its ubiquity, and its influence on moral character. Despite having written these articles nearly 15 years apart, both authors delivered points and covered topics that remain relevant today. However, that is not to say that I found myself in complete agreement or disagreement. Thus, I will preface my argument with aspects of each article with which I found myself in concurrence.

4.1: Common Ground

Duncum. Duncum's argument, in summary, concludes that violent media—specifically that of the gratuitous kind—is: harmful to one's moral psychology; inducive of fear and anxiety (among other emotions); warranting individual and collective response. Duncum ascribes the harm violent media has on moral character to its induction of fear and anxiety and corruption of our moral psychology. He further argues that this harm is exacerbated by the rising ubiquity of violent media. He ultimately concludes that this increasing cycle of (primarily gratuitous) violent media and moral harm warrants individual and collective responses.

Duncum and I share common ground in the following: there is harm engendered by violent media; violent media can induce feelings of fear and anxiety; there is an increasing presence of violence in media; there is a symbiotic cycle between the harm and presence of violent media. However, Duncum and I do not share common ground in the following: all cases of gratuitous mediated violence are harmful; social and educational interventions are problematic in cases of transgressive mediated violence; gratuitously violent media is the only form of violent media that carries a realistic potential for successful intervention, and whose intervention is actively warranted.

Stoner. Stoner's argument, in summary, concludes that violent media—specifically that of the questionable kind—is not harmful to one's moral psychology; not substantively different in artistic merit from permissible media; inducive of fear and disgust (among other emotions); inviting of *paradoxical pleasure* (as opposed to *vicarious sadism*). Stoner argues against any posited harm being carried by violent media by establishing that there is no accepted, distinguishing moral difference between the graphic violence depicted in both permissible and questionable media. To substantiate his argument, Stoner debunks three major contenders of distinguishing differences based on artistic merit, meaningfulness, and audience orientation. He accomplishes this using the following thought experiments: (1) a low-quality production of a highly esteemed, historically permissible piece of media retains its permissibility despite its lack of artistic merit; (2) a permissible showing of a permissible film is considered in which its meaningful segment is cut from the film, displaying only the gratuitous depictions of violence; (3) both questionable films and permissible films centered around the two following audience orientations are conceptually possible: *paradoxical pleasure* and *vicarious sadism*.

Stoner and I share common ground in the following: not all violent media is guaranteed to be harmful; there is no currently accepted distinguishing difference between permissible and questionable media; and the ARA is to be considered false for as long as there is no established distinguishing difference. However, Stoner and I do not share common ground in the following: fear and disgust would not elicit in those experiencing vicarious sadism, and that a gratuitous, permissible film stripped of its meaningful parts would lack meaningfulness.

4.2: ARGUMENTS

On Duncum. To begin my arguments against Duncum's convictions that we do not share, his conception of the scope of gratuitously violent depictions in media that can cause harm is overzealous: He considers the consumption of gratuitously violent media without social/educational intervention as harmful. I argue that this is not the case because—as explored in Stoner's article—there are ample examples of such media that are not considered harmful (e.g., *Schindler's List*).

Duncum ensures that social/educational intervention is problematic in cases of transgressive violence, citing the oppositional nature present in adolescents. I argue that this assurance is unfounded, as there is no conclusive evidence to suggest that such a nature is universal among adolescents. Furthermore, even if this assurance *were* true, it would be insufficient to definitively show that interventions would remain problematic.

Lastly, Duncum considers gratuitous violence as the only form of mediated violence to warrant and to carry the potential of being beneficial. I argue other categories of mediated violence—specifically that of transgressive—have the potential to benefit from social/educational intervention. Regarding transgressive violence, the potential to benefit can be realized by focusing on the severe repercussions of its ends, such as establishing masculinity, avoiding prosecution, defying authority for its own sake, etc.

On Stoner. To begin my arguments against Stoner's convictions that we do not share, his assertion that fear and disgust would not elicit in those who experience vicarious sadism is generalizing. As he attributes paradoxical pleasure to the pleasure found in the fear and disgust produced by violent media, the pleasure experienced from vicarious sadism can be exacerbated by the fear and disgust of the sadist's own actions. Essentially, sadists can be disgusted by and fearful of their own actions, which then contributes to the pleasure they experience from being the one to inflict these actions. This phenomenon is comparable to how the level of pleasure experienced by those who are drawn to the taboo increases as their actions become more taboo.

In his argument against the contended distinguishing difference based on meaningfulness, Stoner considers *Schindler's List* as an example of a meaningful film with gratuitous depictions of violence. However, he attributes this meaningfulness to what follows these depictions, not to the film itself. He uses this supposed lack of meaningfulness to demonstrate that a film can lack meaningfulness and still retain permissibility under the ARA. I argue that the film—along with films with similar content—still carries meaningfulness through this dissection. Essentially, films that solely depict gratuitous violence can still be meaningful without any explicit establishment of meaning. Examples of such films include: *A Serbian Film*[5]; *Salò, or the 120 Days of Sodom*[6]; *Texas Chainsaw Massacre*[7], etc. As is the case with these films, their meaningfulness comes from the gratuitous depictions of violence; these depictions are the meaningful statements these films are meant to represent. The gratuitously repulsive plot of *A Serbian Film* is itself a means of establishing the meaning of the film: a parody of the politically-correct Serbian films financed by foreigners.[8]

On Violent Media. The ubiquity of mediated violence is not up for discussion; it is a substantive measurement of the presence that violent media has in our culture. As for its influence on moral character, I argue that it only has such an wide and intrinsic influence when the viewer—having no knowledge of how to respond—consumes gratuitous depictions of violence, since: gratuitous violence induces habituation and feelings of fear, anxiety, and disgust (among others); there is no awareness—let alone response—regarding whether these feelings are of substance; there is nothing else to gratuitous violence than the violence itself; and the cycles of habituation and emotion are harmful to one's moral character.

For the sake of clarity, my argument regarding gratuitous violence is outlined below:

(P1) Consumption of gratuitous depictions of violence induces fear, anxiety, and disgust, and habituation

(P2) Having no knowledge of how to respond appropriately maintains the fear, anxiety, disgust, and habituation without recourse.

(P3) Fear, anxiety, disgust, and habituation are harmful to moral character.

∴ Therefore, the consumption of gratuitous depictions of violence without knowledge of how to respond appropriately is—without recourse—harmful to moral character.

CONCLUSION

The aim of this paper was to give a comprehensive overview of the contemporary discussion on the ubiquity of violent media and its influence on character. Understanding the nature of violent media's influence on our moral character is imperative to preventing unnecessary censorship of powerful works of art; also, in preventing the occurrence of violent media's consequential instances of real-world violence. Furthermore, developing our understanding of how violent media affects moral character can help us learn when and how to consume harmful examples of such media safely. Two articles by philosophers from opposite sides of the discourse were covered: "Attractions to Violence" and "A Defence of Gorey Fictions". The former, written by Paul Duncum, discussed his ambivalent—yet ultimately supportive—approach to the moral concern surrounding the harmful and expansive nature of mediated violence, its nuances, its prevalence in popular culture, and how we can effectively respond, both as individuals and as a collective. The latter, written by Ian Stoner, discussed his issues with the argument that (seemingly) selective depictions of mediated violence were harmful, expansive, and warranted moral concern. These articles provided a thorough account of relevant discourse, sufficient for an informed disquisition on mediated violence, and thus conclusively achieved the aim of this paper.

End Notes

1. *Food, Inc.* (2008) is an example of media that features violence. I chose to reference the film because it is an example of permissible, violent media. It is known for its candid, graphic depictions of the food industry.

2. *Texas Chainsaw Massacre* (1974) is an example of media that features violence *and* has gained notoriety as a result of these depictions. It is known for its graphic, dehumanizing depictions of mutilation and murder.

3. Research article cited by Duncum in this claim: Spark & Spark 2002, “Exploring the Attractions of Violent Entertainment”.

4. Multiple research articles cited by Duncum in this claim: Cantor 2002, “Media Violence Effects and Interventions.”; Horsley 2002, “Action-Revenge Films”; Zillman 1988, “The Psychology of the Appeal of Portrayals of Violence”.

5. *A Serbian Film* (2010) is a film considered offensive due to its disgusting plot and graphic content; The film features themes regarding the worst of humanity.

6. *Salò, or the 120 Days of Sodom* (1975) is a notoriously graphic film whose gorey depictions of violence serve as a commentary on multiple themes, such as Totalitarianism, Capitalism, Political corruption, etc. The film explores the exploitation of 18 teenagers at the hands of 4 powerful libertines during the time of Fascist Italy.

7. *Texas Chainsaw Massacre* (1974) uses its aforementioned graphic content as a means to expose human audiences to the violence against animals in the meat industry. (Wickman 2013).

8. The film’s writer has also attested that the title of *A Serbian Film* was intended to serve as a metaphor of Serbia’s image at the time: chaotic (Grgić 2009).

Bibliography

Cantor, Joanne. (2002) “Media Violence Effects and Interventions: The Roles of Communication and Emotion,” in: J. Bryant, D. Roskos-Ewoldson, and J. Cantor (Eds), *Communication and Emotion: Essays in Honor of Dolf Zillman*, pp. 197-219 (Mahwah, NJ: Lawrence Erlbaum)

Duncum, Paul. (2006) “Attractions to Violence,” *Journal of Aesthetic Education 40*(4), pp. 21-38

Grgić, Velimir (2009). "'Ovdje će te cijeniti jedino kad crkneš'". *tportal.hr* (in Croatian)

Horsley, Jake. (2002) “Action-Revenge Films,” in: J.D. Torr (Ed), Violence in Film and Television, p. 43 (San Diego, CA: Greenhaven Press)

Sparks, Glen G. and Sparks, Cheri W. (2002) “Exploring the Attractions of Violent Entertainment,” in: James D. Torr (Ed), *Violence in Film and Television*, pp. 114-126 (San Diego, CA: Greenhaven Press)

Stoner, Ian. (2020) “A Defence of Gorey Fictions,” *Journal of Applied Philosophy 37*(4), pp. 511-527

Wickman, Forrest. (2013) "The Ultimate Pro-Vegetarian Film Is the Last Movie You'd Expect," *Slate*

Zillman, Dolf. (1988) “The Psychology of the Appeal of Portrayals of Violence,” in: J. Goldstein (Ed), *Why We Watch: The Attractions of Violent Entertainment*, p. 182 (New York: Oxford University Press)

PUBLISH OR PERISH

Victor Dominguez

Suppose your favorite musician is outed as a raging racist. Clearly their reputation is ruined but how does this affect the quality of their music if at all? Can we still listen to the music guilt free or are the art and artist so intertwined that it is futile to attempt separation? This project seeks to explore this tension, with Martin Heidegger and his works as a focus. The core issue at hand is thus: we often engage in and or are fascinated by works, philosophical texts, films, music, art, etc. that are produced by problematic individuals. These individuals produce impactful works but hold and/or perpetuate harmful views and ideology. In the case of Martin Heidegger's *Being and Time,* it became an influential work prior to his widely known association with the Nazi party. The project at hand thus explores the tension universally found across all disciplines and forms of art. Our favorite works are sometimes made by bad people, but where does that leave our favorite works? During our exploration of this tension, we shall consider various key factors including, moral consideration, moral compromise, our moral threshold and ultimately our urge to separate art and artist. The goal is not to establish an objective methodology to use when faced with works you love but people you don't. Rather, the goal is to gain traction on a path that leads us to critical consideration rather than superficial compromises. When all is said and done, how much are we willing to tolerate from the greats so far as they produce great works? Does the production of great works grant equally great tolerance?

As we proceed, I suggest these concepts are kept in mind as to better grasp our path. Firstly, by moral consideration I mean considering the moral factors that play into the work one is engaged with. For example, your favorite rapper has released a new album that has reinvigorated the genre but has recently perpetuated racist ideology during the press run. You begin to morally consider the implications of engaging with the work because of its association with the artist who in turn has associated themselves with problematic ideology. Once said artist has brought attention to themselves, they have taken it away from the work. Though the work may not have anything inherently to do with the problematic ideology, by association it is also put under the microscope. In a rather strange way, art is enjoyed the most when the artist is invisible. Paradoxically we have favorite artists, but they are favorite insofar as they produce the works we enjoy. Once the connection between art and artist is brought forth by the faults of the artist we are led to morally consider what it means to engage with the art. As it will be observed, separation of artist and art is never a clean cut and often leaves only blurred lines. Of course, this is pointless if not thought about critically (as is often the case) but I argue that ignorance is not a position worth defending. Neither is the idea that if we morally considered everything we would enjoy nothing. I do not disagree completely, it would be extremely inefficient and constantly lead to dead-ends, yet I argue that not even attempting to consider moral implications has made us complacent and tolerant of too many of the moral transgressions of those who are undeserving.

Secondly, the concept of moral compromise and as the name suggests, is the compromise we make implicitly when engaging with works from problematic artists. I hold that moral compromises are necessary as history has taught us without our civilizations being completely different. As mentioned above, it is true that if we established a rigid moral criterion of what we should and should not engage with we would be left with nothing. That being said, one morally compromises when they acknowledge the flaws of the artist but engage with the art anyways. This is a moral compromise in that even though it is the art we are interested in; by engaging in it we are complicit in some form of supporting the artist. Does that mean that by listening to the new album one is automatically signaling that they agree with the artist on all their respective moral stances? No, but the line is very thin and blurred. What it does signal however, is that one is apt to excuse moral wrongs for the sake of enjoying the art, regardless. As I will reiterate, though we may try to separate the art and the artist so long as valuable art is present, we toe a fine line (a tightrope so to speak) where we establish the precedent that artists can be problematic without concern so long as their work is engaged with or admired.

Which brings us to our next concept, the moral threshold. Though we morally consider and compromise, I seek to explore the ambiguous moral threshold, that is the breaking point of our collective tolerance. I consider what an artist would have to do or say to get them to be completely disqualified from any consideration whatsoever. Heidegger's Nazi association was not enough to remove his place in 20th century philosophy. Seemingly our willingness to tolerate is proportional to the perceived values of the works.

We now turn to our focus, Heidegger. When first introduced to *BT*, it felt like an instant click with someone you just met. I understood the gist of what Heidegger was saying (once his terminology was explained of course). In *BT* I found a text that explored and fleshed out philosophical questions and interests I had, that I could not articulate before. As I got deeper into the text I figured I should learn about the author as well and not only his work. It is important to note that I encountered *BT* first before I encountered Heidegger. It is often the case that one will engage with the work before engaging with the producer and that is relevant for the sake of this project. As such, I was disappointed to find that the author of *BT* was a proud card-carrying Nazi. Not just sign on the lawn but active endorsement and association with the Nazi party. There is nuance to his membership but nuance or not he was still a Nazi. It has been an unfortunate part of working on this project that I found how deeply Heidegger's Nazi association ran. At first, there is the initial shock of finding out he's a Nazi. However, that categorization is only the tip of the iceberg. It is during the research that I observed two subtleties brought forth by this discovery, namely: There is an implicit recognition in the value of his work and thus a search for reasons to morally compromise. Insofar as the perceived value of the *BT* is recognized/upheld it becomes a challenge to simultaneously acknowledge the value of the work and the flaws of the man. I invite my readers to imagine themselves on a tightrope balancing a pole between two skyscrapers. It seems that pros and cons add weight to either side, but the goal is not to fall. Heidegger being handpicked by Hitler is a strong wind that is bound to force some to fall off. Is the value of the work so great that it anchors us in the face of such strong winds? My hesitant response is yes; the value does anchor but also sets a dangerous precedent. If we are to consider the value of works in comparison to the moral transgressions of the producers of works, we run the risk of quantifying morality. A utilitarian scale here cannot assist and should not be looked at as a solution. If it appears as if there is no other possibility, then that is my fault. Though take into consideration the value of work, I argue that falling back on a consequentialist position leaves us no better off than our starting position. Should we make any progress in the puzzle it will be recognized by the lack of consequentialist thinking.

My research led me to two poignant articles, namely, *The Political Myths of Martin Heidegger* by Jeff Love and Michael Meng and *Heidegger's Black Notebooks* by Lillian Alweiss. In both works the tension between artist and art is found. Neither article is directed at giving a definitive stance on whether or not to continue reading Heidegger's work, but both works do provide critical points that deserve consideration. In the former work, *Political Myths of Martin Heidegger*, the authors explore the nuance of Heidegger's Nazi association. Early on they state, "He saw in Nazism an opportunity to awaken science from its slumberous state as a decadent form of practical, technical, and purposive knowledge by arguing a spiritual world of philosophical uncertainty and openness in which the question of being might be raised" (Love, Meng 2015, p.45). In other words, for Heidegger, Nazism is a means and not an end, but does this make any difference? At hand we observe the attempt of justification by the means of the consequence. This may sound fancy, but you've heard this a million times before: I will do something bad but something good will come from it. That Heidegger would align with Nazism as a means and not an end does little to nothing in clearing his association. But as it has been emphasized his version of Nazism was nuanced, in it he thought "his Nazism was more truthful to the "inner greatness" of the movement than the technical, dogmatic, and tedious form of Nazism "peddled about" at the time" (Love, Meng 2015, p.46). His Nazism was different or so he wished to believe. Heidegger stated that "his attraction to Nazism came from an interest in reforming German academic life and emphasized that he had quickly become disappointed with the movement's deviation from its original promise as a redemptive force for Germany and Western civilization" (Love, Meng 2015, p.46). Behind all these similar statements we are led to believe that Heidegger's intentions were that of man concerned for the sake of humanity. Seemingly *being* has lost its way and Nazism (out of all ideologies) would be the one to place *being* on the right path. The level of willingness to ignore clear dangers and harms of Nazism for the sake of philosophical actualization is concerning. Evil, even if it's a slightly distinct version of evil, can never be anything but evil.

Interestingly, as the article points out, Heidegger was not one to involve himself with politics. It should also be noted that across his works Heidegger has no ethics. In comparison to Plato they state, "Like Plato, Heidegger shunned the life of the polis, the world of doxa, for most of his life except a brief moment when the tyrant seduced him into it" (Love, Meng 2015, p.47). He was never politically vocal until the worst political ideology *seduced* him. I am critical of this, as it suggests that Heidegger held problematic views long before his seduction. Nazism was not introduced to him like a PowerPoint presentation after which he suddenly decided to join.

Though the article goes on to elaborate further, it is best we turn our attention to the other work, *Heidegger's Black Notebooks*. In contrast to the previous works, here we are less focused on the politics and spotlight his antisemitism. The black notebooks are a series of personal journals of Martin Heidegger published after his death. In the publishing of these notebooks, it became apparent that Heidegger was an antisemite and that in their publication suggested "Heidegger's philosophy was inherently antisemitic" (Alweiss 2015, p.305). This suggestion further complicates our considerations. As stated above, it is clear that the man was corrupted but his works seemingly were not. With the publication of the black notebooks, we now must also confront the possibility that there are antisemitic undertones to be found in his work. It was one thing to separate art and artist, but it becomes an almost impossible challenge to separate book and passage. It seems that Heidegger's Nazism made it into later works and lectures culminating with the view that the Jewish community was the embodiment of his critique on modernity. Basically, they embodied everything he thought was wrong with the world. Alweiss explains, "According to Heidegger, mankind can only be saved if we are able to overcome the age of machination and our will to power and instead become guardians of Being. The Jewish people as a 'race' are inherently prevented from achieving this" (Alweiss 2015, p. 313). Nazi or not, he was still a racist. He even goes so far as to claim "Husserl's thought or philosophy is limited by the fact that he is a Jew" (Alweiss 2015, p.313). It is also important to mention that Heidegger was elected Rector of Freiburg University in April 1933. He joined the Nazi Party on May 1st. As Rector, Heidegger was responsible for the "alignment" or "synchronization" (*Gleichschaltung*) of the university with the aims of the Nazi Party. In April 1934, a scant year later, Heidegger resigned the rectorship, although he remained a member of the Nazi party until the end of the Nazi regime (Zalta, 2025). Heidegger obtains a key position, ousting his Jewish mentor Husserl and thus begins to align the university with the aims of the Nazi party. Heidegger's character is on full display here, but caution is advised not to fall into fallacious reasoning. Like consequentialist thinking, to reject Heidegger's ideas on this basis is to commit the ad-hominem fallacy. We cannot simply reject the work/ideas on the basis of personal attacks on Heidegger. We must strive for a more concrete criterion lest we be accused of reactionary reasoning. Again, we return the imagery of the tightrope balancing act, as clearly it is the personal actions of Heidegger that cause tension. Though, it may seem like no progress can be made, it is a balancing act with out a net to fall on here. If we don't consider the man at all then we are culpable of active ignorance. We cannot simply rely on ad-hominem as justifications not to read or engage but we also cannot ignore his actions. Heidegger never apologized and that matters even if we cannot quite pin down how that affects his work. The list of individual wrongs can go on and on, but I feel the point has been made clear enough.

Forever a stain on his reputation, his association cast a shadow over his work and life up to this day. Though, I am disillusioned and at a crossroads, *BT* was still an important work to me and though Heidegger was a Nazi and there was seemingly no correlation between his personal politics and the text. I have never formally proclaimed this to be the case I have quietly attempted to separate the art and the artist. I distinguished between Martin Heidegger the person and Martin Heidegger the philosopher and kept the philosopher. This seems to be the ideal solution with all the flaws of idealization. Realistically speaking, there is no way to truly separate the art and the artist.

Interestingly, this leads to reflection not just on what has been discussed but on internal reflection of ourselves collectively. The fact that moral compromise is the norm and that artists across entertainment and academic disciplines are full of problematic people who remain popular is the precedent I keep referring to. We have become so obsessed with that we forget there are people who produce them. In a world where social disconnection is increasingly common, we do not consider the people behind the products we engage with. Take, for example, the door dash delivery person who really has no name and your interest is only in the service they provide. Why would anyone consider the moral character of someone dropping off some food? They don't unless they are forced to and that is the point. The person only becomes visible when they call attention to themselves as a person and not the service or product they produce or provide. Suppose Heidegger's philosophy is the product, one can adamantly hold the position of indifference and condemnation for Heidegger and all the reverence they can muster for *BT*. It is in the consideration of the person behind the product that conflict arises because we are so used to only caring about the product, only caring about its use and potential uses. Earlier, I mentioned how consequentialist thinking cannot and will not be the solution to this dilemma. In order to escape falling into consequentialist justification we must take another step back. Before moral consideration we should attempt to reestablish the connection between producer and production. If we only ever consider the producer when they make themselves visible, we fundamentally misunderstand their relationship. The producer (as obvious as it seems) is always behind the product, but our form of understanding leads us to believe there is a clear cut between the two. There is a definitive blur between the two but it is an error to consider both separately from the other. Above, I mention the riddle of how the moral flaws of the artist clearly tarnish them, but it is unclear as to how it affects the quality of the work. The quality of the work cannot be judged by the personal flaws of the artist, but it is this approach that leads to dead ends. We are not asking the right questions. Instead of asking how the moral flaws of the artist affect the art, we should be asking why our focus on the uses of the art blind us to the artist. If our intention is always to get the most from the product then we shall never confront the artist until we are forced to. Granted, focusing on the art exclusively, we are shielded from the long list of questions we ask when focusing on people.

Knowing all of this, why do we continue to read and give credence to such a bad person? We do so, I hold, for various reasons including, historical temporal distance, cultural positionality and lack of a present threat. It has been almost a century since the time of Heidegger making considerations passive rather than active. Because it is passive, it is also no longer a threat (though embers of Nazism still persist). Time grants both us readers and Heidegger some grace. This in no way means to say that his moral transgressions are lost in time. Rather, time has converted Heidegger's acts into objects. We can no longer actively do anything about his association but analyze it. Truthfully, it is time that makes a world of difference, as admittedly, I would not hold the same position without the historical distance in mind. Were a known racist to publish a *BT* equivalent today, I would denounce both work and author and hold in contempt individuals who still read the work anyways. I admit the possibility of contradiction but maintain that on principle, if we still elect to read work from a problematic source on the grounds that said work is of some value we will have traded in our morality completely for perceived value. This will only embolden those with problematic views to be more open knowing that so long as great works are produced, they are untouchable.

Thus, we arrive at the concern behind the rest, something that has been hinted at throughout the work, namely, the search for a moral threshold. With or without the consideration of time we tolerate to the point of fault and it leaves us no better off than no consideration at all. Across philosophy and all the arts, we consistently find amazing works and the problematic people that produced them. The question then is when it is too much, that is to say, when we have tolerated too much? I ask with Heidegger in mind, as his Nazism was and is tolerated. Even in his era, during and post WWII he was still widely recognized. What I say here is bound to ruffle feathers but there has to be a limit of some kind. We will not stop compromising as moral perfection is unachievable. Thus, we must take the good with the bad. We cannot throw the baby out with the bathwater, but I ask you to consider how dirty the bathwater is. Clearly, Nazi association is not grounds for complete disqualification. Typing that feels insane, as I would not be friends with someone who was a Nazi. But then temporality becomes a factor and we have to consider when they were a Nazi and if they are still a Nazi. That is for another work, but it remains the task to find out what position, regardless of temporality, cannot be ignored. What position can time not make an object? Aristotle justified slavery but is widely known for his book on ethics. Would we now collectively allow for someone who justifies slavery to also speak on ethics? This is not a call for retroactive moral judgment on the greats, we are not trying to *cancel* Heidegger, Aristotle, and/or the entirety of the western practice of philosophy. It seems though that we should be expected to have learned from the mistakes of the past and as such apply critical criteria for future work. We cannot passively continue to allow admittedly great artists to feel free of consequence so long as their work is perceived as valuable.

As I have emphasized above, were we to consider all moral factors we would never enjoy anything or engage with anything. As of now, our position on the tightrope is still very far from the other building. Regardless of the difficulty, we should consider, as much humanly possible, who the artists we engage with are and what our engagement perpetuates.

Bibliography

Alweiss, Lilian. “Heidegger’s Black Notebooks.” *Philosophy*, vol. 90, no. 352, 2015, pp. 305–16. *JSTOR*, http://www.jstor.org/stable/44015160. Accessed 27 Mar. 2026.

Love, Jeff, and Michael Meng. “The Political Myths of Martin Heidegger.” *New German Critique*, no. 124, 2015, pp. 45–66. *JSTOR*, http://www.jstor.org/stable/43910657. Accessed 27 Mar. 2026.

Wrathall, Mark, "Martin Heidegger", *The Stanford Encyclopedia of Philosophy* (Spring 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/spr2025/entries/heidegger/>.

PLACING WHERE AVERAGENESS BELONGS

Jonathan Lima

INTRODUCTION

Generally, individuals possess different degrees of ability (some innate or cultivated), in quality and quantity. For specific fields, certain abilities are required for qualified or specialized participation. Some may even be exceptional at the field. So, people can be separated into camps of qualified specialists or ordinary outsiders. This paper examines the normative claims about the relationship, meaning a claim of what ought to or should be the case. The question is: should various specialized institutions focus on the interests of the public, or should ordinary individuals exert more effort to understand the work done in specialized institutions? For transparency, I am not convinced of either position.

Of course, the question may be applied broadly to any field or aspect of life. Also, one may point out that within each of us sits some activity we are particularly *good* at. There may exist a field we excel at, as we possess ordinary and specialized (or exceptional) abilities. To limit the scope, we will only consider the field of academic philosophy. The choice is semi-arbitrary, but philosophy can easily lend itself to this debate, which I will explain later. The sections that follow will be: first, setting up the concepts; second, the articles to be used and their content; third, the relationship between the ordinary and specialized in philosophy; and fourth, the conclusion.

SECTION I: CONCEPTS

The usage of the term field will mean "academic philosophy," unless stated otherwise within the sentence. Ordinary (layman or average) identifies individuals that have common qualities in ability and lack professional status with respect to the field. To be identified as ordinary requires comparing the individual with the requirements of the field (e.g. job requirements) or with those of individuals specialized in the field. It may be externally imposed, such as failing to meet the requirements to enter the field or labeled 'ordinary' by being an outsider to the field. Alternatively, it may be internally imposed, with the individual intending to remain outside the field or self-identifying as ordinary within it.

‘Specialist’ identifies an individual possessing the ability to meet the standards for professional participation and who are formally recognized as qualified practitioners in the field. ‘Identification’ measures the individual's ability to meet the field's standards. However, there is a greater emphasis on external imposition, as specialization requires institutional acceptance. While internal impositions may exist, they still depend on some external validation to actualize the identification as a specialist.

‘Exceptional’ identifies individuals with abilities far above the required standards in the field. These exceptional individuals could be considered to have demonstrated mastery of the field or gained social notoriety as a voice of authority for a given field. It does not necessarily require external imposition to validate identification, as one can be ordinary and exceptional through talent that remains unknown or unrecognized (intentional or otherwise). So, exceptionality can be either a latent quality or a social label. But external validation does provide some level of formal recognition and usually comes with being identified as specialized. Of the three concepts, exceptionality will not be discussed in detail. It is defined only for the purpose of acknowledgement.

Based on these definitions, we will establish two groups for the rest of the discussion: the general public (comprised of ordinary and, occasionally, exceptional individuals) and the specialized in-group (comprised of institutional practitioners, whether exceptional or not). The following pieces provide different perspectives on their relationship.

SECTION II: CONTEXT

Now, I borrow the idea (which is not original in itself) for this division from David Hume's *An Enquiry Concerning Human Understanding*. The *Enquiry* established his claims of empiricism (a view that all knowledge comes from the senses) and skepticism (the general doubt toward claims of truth), but the first section provided clarity for the definitions I had previously created. The first section, "Of the Different Species of Philosophy," criticizes the *abstruse* philosophy in favor of the *easy* philosophy. The paper defines abstruse philosophy as being highly theoretical, abstract, and done in isolation. In contrast, easy philosophy leans into humanistic questions and concerns. To defend the abstruse philosophy, Hume acknowledges that it engages in clarity and precision and the existence of skills which are used to further enrich philosophy. Here, too, he discusses the relationship between the “generality of mankind” and their consumption of philosophy, so that it is understood why they choose the easy philosophy. The easy philosophy is not necessarily easy because it's low brow, but because it has wider applicability. It can be used in everyday life, Hume noting, the easy philosophy "moulds the heart and affections," (Hume 1748, pp. 2) making it relatable.

The issue of my paper is further alluded to when Hume states, "by touching those principles which actuate men, reforms their conduct, and brings them nearer to that model of perfection which it describes" (Hume 1748, pp. 2). As a very general understanding of philosophy, which ordinary people most likely possess, it is valuable for elevating and cultivating *better* character based on different principles (e.g. virtues). So, if the ordinary can meet the principles proposed by the easy philosophy, could they not also meet those of abstruse philosophy? The question is not directly answered, but my answer would be yes. Many specialists in philosophy begin as laymen, but through dedication to the craft meet the standards of abstruse philosophy.

Still, Hume and other philosophers have been discussing philosophy's relationship with the public since antiquity. Modern discussions of philosophy are no different, with various arguments being made. So, I will discuss three articles with themes that can be extrapolated from to fit the context of this paper.

The first of the three articles is by Stijn Conix, Olivier Lemeire, and Pei-Shan Chi, titled *The Public Relevance of Philosophy*. Of the three articles it is the densest as it makes use of empirical data to answer their philosophic questions. The researchers argue for the normative position that because philosophy is at times publicly funded, they should involve themselves with public interests. There is no issue with the claim. Any funds made by the work of others should consider their needs met first.

The authors ask how often academic philosophy maintains public relevance. To investigate, they separate studies in philosophy into "philosophy that engages with other fields of science... *Philosophy of Science* (PoS)... philosophy involved with identifying and understanding what we value... *Value Theory* (VT)... LEMM—*philosophy of language, epistemology, philosophy of mind, and metaphysics*" (Conix, Lemeire, Chi 2022, pp. 3). LEMM is assumed to be the most isolated from the public because the content is relevant only to the field itself. It would be very difficult to expect the ordinary, who lacks background knowledge of the subject, to engage with LEMM. Compared to PoS and VT, which are easily applicable across disciplines and discusses issues that the public may want to hear more about.

To verify an answer they distinguish between two types of isolation: *content isolation* and *uptake isolation*. The former is isolated because it is relevant only to those studying the type of philosophy. The latter is isolated by the measure of how often ordinary individuals or other disciplines make references to the subject. To investigate the claims, they used "two empirical tests—one using alt metric data of philosophical journal papers, and one using the abstracts of philosophical journal papers" (Conix, Lemeire, Chi 2022, pp. 3). Alt-metrics are the engagement with and mentions of a philosophical paper on non-academic platforms, like social media (test two, uptake isolation). The use of abstracts is to evaluate if the paper is on an isolated topic (test one, content isolation).

The findings confirmed many assumptions. First, test one establishes that most academic philosophy topics are content isolated. Test two supported the claim, as alt-metric data was low. Next, test two supports the initial assumption that LEMM shows more 'uptake' isolation than PoS and VT. VT was the least uptake isolated of the philosophies. Test one continues to validate VT as the least content isolated. However, there was no difference between PoS and LEMM in content isolation. So, PoS and LEMM only differed in uptake isolation as previously mentioned.

All the findings were somewhat straightforward. I assume, LEMM is too theoretical for anyone outside the field to meaningfully consume the research. PoS is also as theoretical, but due to being political, I imagine would not be something ordinary individuals want to discuss often. VT intuitively makes sense to be more engaging for the public. The public enjoys conversations on content that experience daily, such as questions of morality. Although all three still did not have large amounts of alt-metric data to analyze, which could mean the public simply is not interested in philosophy's offerings.

With these findings, the authors shift to asking why many of the works are isolated in the field. They answer that incentives in the institutions of philosophy are structured to promote isolated work. Specialists are under constant review, consideration for promotion and tenure (RPT), with their work reflecting their worthiness of these gains. Also, funding is allocated to research that furthers the philosophy the specialist is in. Thus, the conclusion is to change the structure of academic philosophy to incentivize work in public philosophy.

The second article is by William G. Lycan, titled "On the Uses of Philosophy." Lycan uses the paper to respond to Philip Kitcher's critique of modern philosophy. The following is a broad synthesis of Lycan's article. Lycan agrees with many of Kitcher's views of philosophy. Such as

Kitcher's stance that "the professionalization and the increasing specialization of philosophy" (Lycan 2024, pp. 548) has negatively impacted the field, and claiming the work done now in these isolated fields is of interests only to themselves. Lycan also discusses the importance of applied ethics and public philosophy such that there should be greater emphasis on both.

Lycan takes issue with the final implied normative claim. That individuals should pursue applied ethics and public philosophy compared to specialized philosophy. Lycan is not against the former philosophic fields, as he agrees there should be more. Instead, "pure" or specialized philosophy should not be abandoned when work in either is not mutually exclusive. He tries to understand the "should" in Kitcher's claim. Being a utilitarian (a theory that the best and most moral action is one that maximizes happiness for the greatest number), there is some pull towards doing philosophy that many can utilize. But specialized philosophy is utilized by other fields and is done out of intrinsic value. Similarly to Hume, he defends specialized philosophy as fundamental to general philosophy. Also, specialized philosophers do teach the skills necessary for their field to those not in philosophy. As a conclusion, Lycan discusses topics where specialized philosophy is a direct part of the conversation, like consciousness, and for that reason should not be disregarded.

Yet, Lycan's conception of curiosity, which is the first point he makes, has informed much of this paper. He states that humans have broad "natural inclinations," shaped by internal and external factors, that spark curiosity in us. These feelings guide us towards specific pursuits. In the case of pure philosophy, a certain type of curiosity is needed to see its value. Lycan possesses such a disposition, hence his enjoyment and pursuit of the field. Of course, individuals will or won't share the same interests. Lycan explains, "that distinctive curiosity is not for everyone... the world might be considerably worse off if very many people had it... That's fine. It is as it should be. Intellectual diversity is vital to society; in fact, it's vital to academia" (Lycan 2024, pp. 549). The phrase, "it is as it should be," informs me of the interpretive conclusion. To dictate to others what they should be interested in goes against an individual's naturally provided agency.

The last piece is Meghan Sullivan's review essay *Philosophy as a Way of Life*. Sullivan sets up the discussion by contrasting the mid-twentieth-century view of moral philosophy with current discussion. Consider mid-twentieth-century moral philosophers as claiming that philosophers have little authority about moral claims. Sullivan explains the history for why these philosophers came to such a conclusion. She cites C.D Broad's statement that "moral philosophers... have no special information not available to the general public about what is right and what is wrong" (Sullivan 2023, pp. 588). Any individual could go and become more well informed on the topic. With the public and the moral philosopher sharing the same pool of information, neither has better claim for an authoritative position. The only differentiating qualities is the conceptual analysis applied to moral questions by moral philosophers.

Yet, twenty-first century discussions, led by a mix of several specialized (and non-specialized) individuals, have seen the re-emergence of life advice focused moral claims, like the virtue focused philosophy of the past. Unlike applied ethics, which measures the morality of an action, this applied tradition philosophy is focused on normative claims. These authors create pieces that state what the reader's "life goals" should be and why. Sullivan goes on to review the *Guides to the Good Life* book series, published by Oxford University Press, to examine the quality of the claims.

Naturally, Sullivan finds many different proposals of what the *good life* could be. The distinction between argument and proposals is that the former is strict in its evaluation. Hence, applied ethics makes use of methodological argumentation because it is concerned with the universal applicability of a moral claim. The arguments must be convincing based on their logic. However, the latter is giving the reader recommendations and desires that they emphasize agency in accepting certain claims. Clarity is still a requirement, but its use is so that the reader can genuinely understand the overall life the author claims they should live. The dynamic between the author and reader is similar to "friends or life coaches" (Sullivan 2023, pp. 607). Where intellectual hospitality is provided to newcomers and information is more readily democratized.

Still, applied tradition is not free of controversy. There are concerns about individuals gaining a pedestal within the specific applied tradition. Or the worry of perpetuating flawed and harmful claims simply because practitioners may desire faithfulness to the tradition. Also, despite all the democratization, there are some traditions that require time to even begin to understand. Such a task would be no different from demanding that the layman begin an academic journey simply to understand a specialized philosophical concept. Thus, the applied tradition is not necessarily a better approach to public philosophy, only a different one.

Ultimately, each paper explains what philosophy means to those outside of it. The first article is a study of why philosophy is not more engaged with the public. The second article I consider explaining the overwhelming presence of specialization in the field to be a product of choice. Lastly, the third article engages with an approach that desires to directly involve the ordinary. Each is positioned from the philosopher's perspective. What would be the ordinary perspective?

SECTION III: THE ORDINARY AND THE SPECIALIST

Now, when those in the ordinary position communicate with a specialist in the field, without requesting layman's explanations, they will quickly find themselves confronted by jargon. Jargon is terms without context or are difficult to understand for the layman. For some, after the confusion, they would consider the person explaining these concepts to have great abilities. To the ordinary, they are exceptional. Others would find these unknown concepts so strange that they are inapplicable to reality. Of course, for a subject like philosophy, many of the claims can be applicable to their lives, but the ordinary person does not know better. Also, if they knew better, the ordinary person may not consider the specialist exceptional, just as the specialist does not. So, should various specialized institutions focus on the interests of the public? On the small scale, it would mean the philosopher in the example should have begun with very general explanations. On a larger scale, academic philosophy would have to lower the barrier of entry. It could be through lowering the applicant requirement in the field, but that would be deemed either extreme or an unrealistic demand. Instead, as both Lycan and the first article state, the conversation in philosophy should be broadened.

But would this change the public perspective of philosophy? Parallel to the incentive structure described in article one, ordinary living is comprised of systems that do not incentivize philosophic thinking. The ordinary is, for the most part, in a state of work. When they do have time for leisure, most would prefer indulging in relaxing activities. Not that philosophy could not be relaxing, but for common folk, it demands some attention to grasp. Additionally, ordinary living does not always reward critical thought. Instead, it often rewards the passive individual, who nods their head, shows up to work, and makes no fuss. Many of us can account for moments or large periods of routine where our minds could be considered *shut off*. Philosophy does not allow the individual to lay down their mental abilities. Lastly, other systems of society contribute to general negative biases towards philosophy. If philosophy began to speak on topics from their ordinary lives, it would not guarantee they would pay attention to the claims. Thus, I doubt appealing to a group that already dismisses the field will attract engaged attention from them.

However, another way to include the public is to make philosophy more literarily accessible. Examples include more open (i.e. free) journal publications, writing, or speaking of specialized research for a general audience. The data of article one does showcase the existence of a public audience for philosophy. Even if the engagement could be passive, an example would be an ordinary person listening to pop philosophy from podcasts. They may very well enjoy the claims, but once the piece of media is put away, so too is philosophic thinking. Still, accessibility for a handful is rarely morally objectionable. The more it is readily available, the more likely people are to engage with the field. So yes, institutions should focus on the interests of the general audience if it is for the purpose of allowing easier access to philosophy.

Alternatively, if the ordinary person shows interest in philosophy, would the obligation of effort not fall on them? Simply because an individual desires to participate does not change the requirements in place. Think back to Lycan's curiosity concept. The agency, from curiosity, affords you the disposition to desire a given field. The requirements should be accepted as they exist to make sure everyone is on the same page. Philosophy should not have to change to meet the ordinary person at their level. So, the ordinary would have to read the foundational material, conversations with professional academics, or so on. Since they found the field interesting enough, they may continue diving into the field's particulars. Also, those in the field did not get to their position without considering and meeting the requirements. In that case, why could the ordinary not meet them as well? The individual is responsible for their choice in engagement, not the party with. However, if the specialized philosophy is being elitist, then that would make the field responsible, as it is unfair to the public. So, yes, ordinary individuals should exert more effort to understand the work done in specialized institutions if they want to participate.

Yet, I find the neutrality of claiming both uneventful. Let us reevaluate the incentive claim. For an ordinary individual in this camp, entrenched in the malaise of social norms, they most definitely will be preoccupied with other subjects, and not philosophy. Now, it is of no fault of their own. Their perspective has been limited to what is provided by the system. The innate abilities they may have are never realized because the system does not incentivize them. The only guide most ordinary individuals have is their passions. Rational or critical philosophical thinking requires mature forms of conceptualized analysis. Maturity is a product of sustained experience with dense concepts. But, for the public, who often do not engage with these subjects, they possess none of this maturity. What they do have is curiosity or passions that guide them towards specific interests.

For example, the vague assumptions and ill-informed understandings of a subject one may have are guided by intuitions. You make a guess in the dark for what it is you believe the subject could be, with very little experience of it. Experience could be direct or indirect. Indirect experience of philosophy would be through entertainment media and direct experience from seminars. Either way, you have a vague acquaintance with the subject. I consider life to be one of these vague subjects. Ask an ordinary individual about living, and the answers will vary, but the vagueness of the reasons for the answer always remains. Life is a heavy subject, with too many branches to conceptualize. It cannot be understood from only one field and understanding it from as many fields as one can creates the vagueness, like simple conceptions of philosophy. The connection to the normative question from the start is as follows: yes, to both claims. Philosophy should discuss a public topic, and the ordinary individual should exert themselves to participate in philosophy. But it is the applied tradition that I argue should be the public topic.

Argumentative philosophies, easy or specialized, require the ordinary to find the logic of the claim and afterwards be convinced. It is very methodical, and for those untrained in mature thinking, understanding even simple justifications could be grueling. The applied tradition only requires agency, or the curiosity to follow a tradition. Since the public is well accustomed to their passions, the entry barrier is low. Following a tradition forces the ordinary person to internalize new ways of thinking about life and analysis.

Now, Sullivan did make the point that there could be malpractice issues. To avoid these issues, philosophers would have to be more community oriented. Philosophers check one another's works and provide clarity when needed. Most fields of philosophy already do this, but emphasizing critical thinking in this field would have to be a constant necessity. If one tradition, party, or individual gains enough notoriety without challenge, dogmatism sets in. There would also be a question of how research in the field could be conducted, beyond the given advice. Maybe the field would emphasize the public voice.

The field would be similar in form to online groups with public presence. Philosophic influencers providing advice based on researched literature. They would maintain a critical perspective of their own findings. In time, the public might naturally feel more inclined to branch into different philosophical fields. As for those who were guided only by passion, with enough time and in the presence of critical traditions, they could change their methods of thinking.

SECTION IV: CONCLUSION

The dynamic between the public and specialized philosophy, or philosophy generally, might never be bridged. Implicit in that statement is the positivity that is found in philosophy. Philosophers are certainly engaging with thoughts that concern them, and those conversations have value. Sure, there might be public interest in more readily available, summarized philosophy, but specialists do not seem interested in those conversations. And why would they be? Those curious about philosophy are already a niche group. Not exceptional, only niche. They chose philosophy most definitely not for the money, but for reasons well beyond public understanding.

Some would disagree, and they have a right to do so. Their curiosities land with the public. If they can convince more of the institution to change the academic infrastructure, then even more power to them. Philosophy should not be stagnant, and it rarely is. Old traditions reemerge, and new ones are just as frequently emptied out.

Thus, this paper serves only to raise questions on the issue. The three articles provide depth of context from different approaches to the field. One is of a sort that desires to prove the need for more public philosophy through empirical testing. The other claims that the specialized field should not be left behind and can be enjoyed alongside the less frequently practiced. The last article points to a promising renewed philosophy that hopes to give advice rather than articulate a logical point. All serve as rebuttals and acceptances for the ordinary sort of individual who may find themselves clamoring for some philosophy.

Bibliography

Conix, Stijn; Lemeire, Olivier; and Chi, Pei-Shan. (2022) "The Public Relevance of Philosophy," *Synthese: An International Journal for Epistemology, Methodology and Philosophy of Science 200*(1), pp. 1 – 28

Hume, David. (1748) An Enquiry Concerning Human Understanding (Indianapolis: Hackett Publishing Company)

Lycan, William G. (2024) "On the Uses of Philosophy," *Philosophia: Philosophical Quarterly of Israel*, *52*(3), pp. 547–557

Sullivan, Meghan. (2023) "Philosophy as a Way of Life.," *Ethics 133*(4), pp. 587–609

CONTRIBUTORS

Rafi Bustos. (Pursuing) BA Philosophy, California State University-Los Angeles. Areas of Interest: Feminist Philosophy, History of Philosophy (specifically 19-20th c.), Deleuze, Aesthetics, Existentialism, and Philosophy and the Emotions. Future Plans: Graduate School (Masters and/or Doctorate) or trade school (automotives).

Victor Dominguez. B.A. Philosophy, Cal State University Dominguez Hills. My areas of interest include, ethics, Heidegger, Nietzsche, the problem of free will, Phenomenology, and hermeneutics. My future plans are to pursue a doctorate in Philosophy.

Giovanni Galvan. B.A. Philosophy: Pre-Law, California State University Los Angeles; B.A. Psychology, California State University Los Angeles. Ethics, Formal Logic, Metaphysics, Philosophy of Science. Giovanni plans to apply to J.D. programs and attend law school.

Daniel Garcia. B.A. Critical Race, Gender, and Sexuality Studies, California State Polytechnic University, Humboldt. Trans Philosophy, Decolonial Philosophy, Philosophy of Self, Metaphilosophy. Daniel will apply to Philosophy Ph.D. programs upon completion of his M.A. degree.

George Getze. B.A. Philosophy, Cal State LA. Areas of Interest: Ethics, Feminist Philosophy, Philosophy of Race. I am planning on pursuing a master's degree in philosophy, potentially followed by a PhD in Philosophy.

Han Kirkegaard. B.A. Biomedical Engineering, University of California, Davis. Trans Philosophy. Han is considering applying to Ph.D. Programs upon completion of their M.A. Degree.

Jonathan Lima. B.A. Philosophy with a minor in History, California State University, Los Angeles. My areas of interest are ontology and applied ethics, mainly to recontextualize the lived experience from a holistic, meta-ironic perspective. As for plans, I am seeking to enter a sedentary occupation to later reflect on my willfulness towards the systems around me.

Niyaz Mahmud, M.A. Philosophy, California State University Los Angeles. M.S. Computer Science, California State University Fullerton. B.F.A. Art, California State University Fullerton. My current interests lie mainly in Metaphysics and Aesthetics. I want to keep the option of pursuing a Ph.D. open, but as an avid general hobbyist, I would like, in my lifetime, to participate in at least one heist and foil at least one heist.

Diego Irigoyen Monteiro. B.A., Arizona State University. I am mostly interested in the philosophy of language and metaphysics My plans for the future are to continue studying philosophy independently.

Ekaterina Vorozhtcova. B.A. Philosophy, California State University Dominguez Hills. My philosophical interests center on metaphysics, philosophy of language, and Continental philosophy, especially questions about the structure of reality, ontological commitment, meaning, and the limits of language. I plan to continue advancing my research and writing in philosophy, with the goal of pursuing doctoral study and contributing original work to the field.

Joseph Yoon. B.A. Stanford University. Meta-ethics, philosophy of action, metaphysics, epistemology. Applying to PhD programs in philosophy.

Master of Arts in Philosophy
California State University, Los Angeles

The Department of Philosophy at California State University, Los Angeles offers a program of study leading to the Master of Arts degree in Philosophy. The program aims at the acquisition of a broad background in philosophy. It is designed for those preparing for further graduate study or community college teaching, and for self-enrichment. Although the department is analytically oriented, it encourages work in other areas, for example Asian philosophy, feminist philosophy, and the interaction between European and Anglo-American thought. The Department includes faculty members with diverse backgrounds and interests actively working in a wide range of philosophical specialties. Classes and seminars are small with a friendly, informal atmosphere that facilitates student-faculty interaction.

The academic programs in philosophy at California State University, Los Angeles are intended to engage students in philosophical inquiry. They aim to acquaint students with noteworthy contributions by philosophers to the tradition; to explore various philosophical issues, problems, and questions; to provide students with principles of inquiry and evaluation relevant to the many areas of human activity, such as science, law, religion, education, government, art, and the humanities; to develop in them skills of analysis, criticism, and synthesis needed for advanced work in various scholarly fields; to encourage the development of skills and attitudes leading to self-reflection and life-long learning.

Philosophy in Practice
Submission Information

Each of the student contributors was specially selected to submit a paper for this issue of *Philosophy in Practice* by one or more faculty members in the Department of Philosophy at California State University, Los Angeles. All writers are currently either students in the master's program of philosophy or undergraduate majors in philosophy. All philosophy students at California State University, Los Angeles are eligible for nomination, and those who were chosen to contribute have demonstrated a superior ability to develop and compose works of advanced philosophical writing.

For more information on *Philosophy in Practice*, please contact:
mshim@calstatela.edu

Department of Philosophy
California State University, Los Angeles
5151 State University Drive
Los Angeles, CA 90032-8114

Phone (323) 343-4180 • Fax (323) 343-4193

www.calstatela.edu/dept/phil