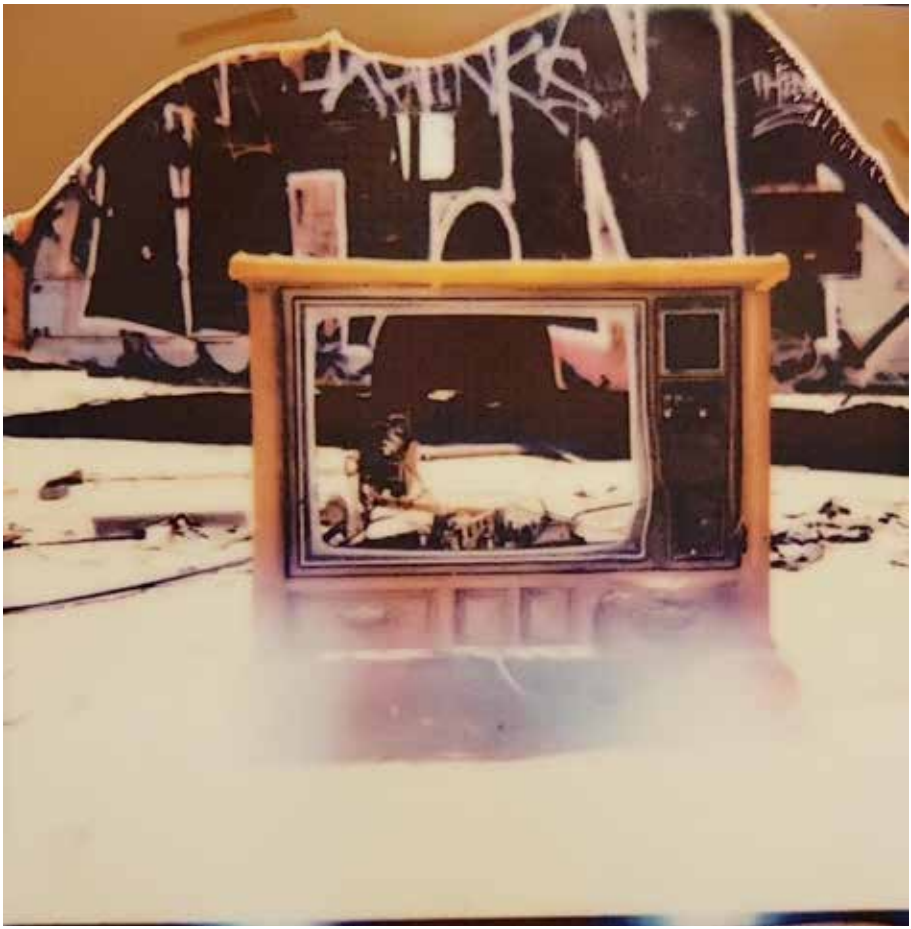
Philosophy in Practice

VOLUME 18 – SPRING 2024



CALIFORNIA STATE UNIVERSITY, LOS ANGELES
DEPARTMENT OF PHILOSOPHY

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. 19th-20th 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

Steven R. Levy (2005-), Ph.D. University of California, Los Angeles. Epistemology, Metaphysics, Logic, Applied Ethics

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. 20th 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

Bound Variables, E-Type Pronouns, and Donkey Anaphora

Christopher Caldwell

Introduction

In linguistics and the philosophy of language, donkey anaphora presents a fascinating puzzle where semantically intuitive, meaningful sentences elude straightforward logical characterizations. Specifically, this term, introduced by Peter Geach (1962), refers to sentences where an anaphoric pronoun appears to pick out its referent from a non-referential, quantified expression, creating a logical conundrum. Geach's solution revolves around treating anaphoric pronouns as variables that covary with their indefinite antecedents. Such treatment suggests that, in all instances of pronominal anaphora, anaphoric pronouns are bound by quantifiers in their antecedents, thereby obtaining their referential meaning.

However, Gareth Evans (1977), drawing inspiration from Saul Kripke's ideas on referent-fixing and *a priori* knowledge of contingent propositions, presents a counterpoint to the bound-variables theory. Evans introduces his descriptivist theory of e-type pronouns, contending that not all anaphoric pronouns that covary with indefinite antecedents are bound variables but rather derive their reference through descriptive means, a stance that significantly diverges from Geach's interpretation.

In this paper, I argue that Evans's e-type theory of anaphora, although an influential and necessary work to demonstrate further problems associated with pronominal anaphora, fails to establish the referential connection between anaphoric pronouns and their indefinite antecedents. In Section I, I outline a few relevant linguistic definitions and discuss the problem of donkey anaphora. In Section II, I argue for a Russellian treatment of indefinite antecedents and provide what I will call Argument Against Anaphoric Pronominal Reference. In Section III, I examine Geach's bound-variables theory of donkey anaphora and Evans's objection to the theory. In Section IV, I explain Evans's treatment of e-type pronouns and his reliance on Kripke's referent-fixing via description and contingent statements knowable *a priori.* Finally, in Section V, I expound on why Evans's treatment of e-type pronouns fails to establish the referential connection it claims to.

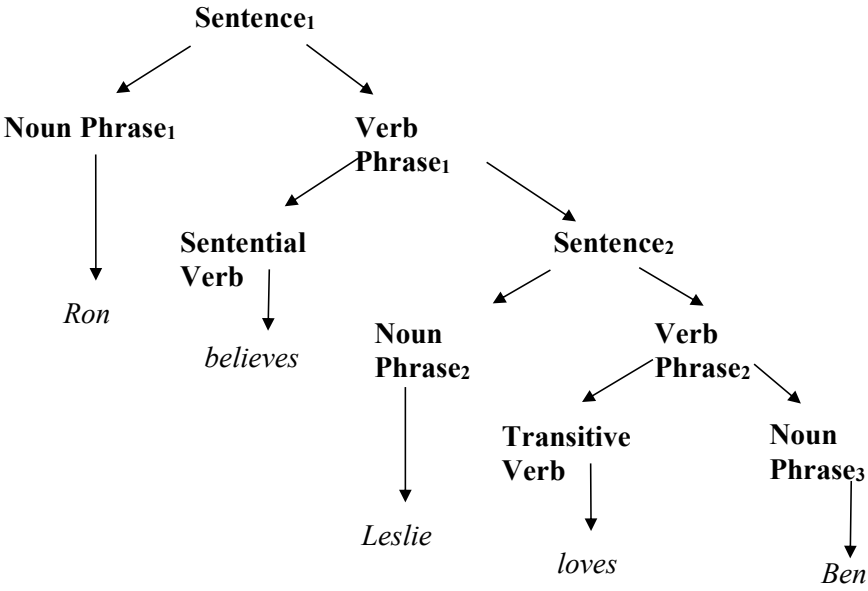
Section 1: The Problem of Donkey Anaphora

In linguistics, *anaphora* is a term used to describe a scenario where the meaning of a linguistic element—such as a phrase, clause, or word—is contingent upon a preceding segment of discourse. This initial segment is often called the *antecedent* of the anaphoric expression. While this description of anaphora could encompass various linguistic phenomena, I focus explicitly on *pronominal anaphora*, in which the dependent element is a pronoun—'he,' 'she,' 'it,' 'they,' and so on. Pronominal anaphora, akin to other forms of anaphora, can occur within the bounds of a single sentence—termed *intra-sentential anaphora*—or it may span multiple sentences, known as *cross-sentential anaphora*.

To set up the dilemma the donkey anaphora poses, it is essential to distinguish between "untroublesome" and "troublesome" anaphora. However, to properly delineate these two types, one must know some technical definitions and be familiar with syntactic tree diagrams.

1.1: Tree Diagrams and Linguistic Terms

In linguistics, a *tree diagram* is a visual representation used to depict the hierarchical structure of sentences according to syntax rules within a particular language; it visually breaks down sentences into their constituent parts, such as phrases and words, showing how these components are organized and relate to each other within the sentence's overall structure. For example, an accompanying tree diagram for 'Ron believes Leslie loves Ben' would look like this:



A *node* refers to a point in the diagram that represents a syntactic unit, such as a word, phrase, or sentence component, indicating its role and relationship within the hierarchical structure of the sentence. The bolded items in the above example denote the nodes of that particular tree diagram[1]. In generative grammar, *domination* is a term used to describe the hierarchical relationship between nodes within a syntactic tree diagram. A node *X* is said to *dominate* a node *Y* if one can trace a path from *X* to *Y* by moving downward through the tree, passing through zero or more intermediate nodes. So, using the previous diagram, 'verb phrase$_1$' dominates, among other nodes, 'noun phrase$_2$' but does not dominate 'noun phrase$_1$'.

The following definition, *c-command*, is critical in an anaphoric context. A node *X c-commands* a node *Y* if and only if all of the following conditions are met:

1. Every node dominating *X* also dominates *Y*.
2. *X* does not dominate *Y*.
3. *X* and *Y* are not the same node.

In other words, c-command is a way to describe how one part of a sentence can govern another part within the sentence's tree structure, without one being directly under the other in the hierarchy. Illustrated by the given tree diagram, 'noun phrase$_1$' c-commands 'verb phrase$_1$' and everything that 'verb phrase$_1$' dominates in virtue of their being dominated by 'sentence$_1$'.

Next, *binding* refers to the relationship between pronouns (and other types of anaphoric elements) and their referents, which are often nouns or noun phrases. The relationship between binding and c-command is foundational to understanding how theories of anaphora categorize different kinds of pronouns and their referents. Put simply, an anaphor—for present purposes, pronouns—must be *bound* in its local domain, meaning the anaphor and its antecedent must be within the same sentence or clause, and the antecedent must *c-command* the anaphor. In subsections 1.2 and 1.3, I will illustrate binding and c-command in anaphoric contexts in more detail.

Lastly, *covariation*, in the context of binding, c-command, and anaphora within linguistics, refers to the relationship between a variable (often an anaphor or pronoun) and its antecedent, where the interpretation of the variable changes depending on the choice of the antecedent. Covariation is especially relevant in sentences involving quantifiers like 'every,' 'some,' or 'no,' or

in conditional sentences, where the existence or identity of the antecedent might depend on certain conditions or contexts.

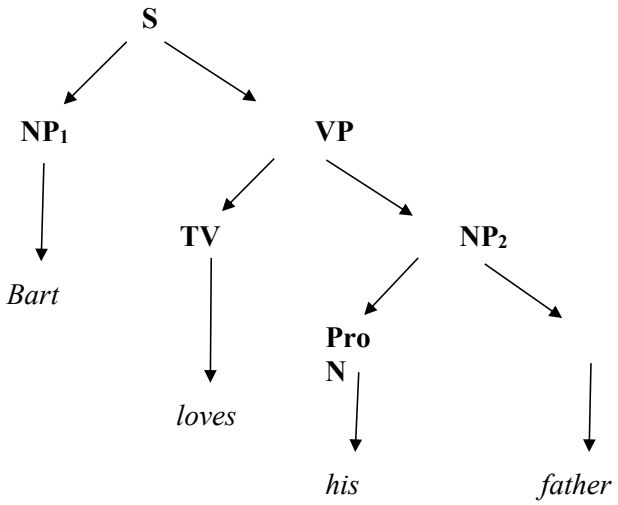
Since instances of covariation involve treating pronouns as variables, symbolic translations in first-order predicate logic are typically used. As such, I will do the same when addressing examples in which the pronoun covaries with its antecedent[2]. As with binding, examples in the following subsections will make covariation clear.

1.2: UNTROUBLESOME ANAPHORA

In most cases of pronominal anaphora, the pronoun receives referential information directly from the antecedent by binding or covariation. This direct reference is especially true for singular anaphoric expressions commonly associated with proper names. For instance, one can easily analyze the following intra-sentential anaphoric statement:

(1) Bart loves his father.

Here is a corresponding tree diagram:



The pronoun ‘his’ picks up its reference from ‘Bart,’ which acts as the antecedent for the anaphoric pronoun. In this case, ‘his’ is bound in virtue of the fact that NP_1 c-commands VP and everything that VP dominates. Following this intuitive line of inference, in singular instances of pronominal

anaphora involving proper names, the referent of the anaphoric pronoun is the same as that of the antecedent (King and Lewis 2021, p. 4).

Cross-sentential examples of pronominal anaphora work the same way regarding proper names. For instance, consider the following example:

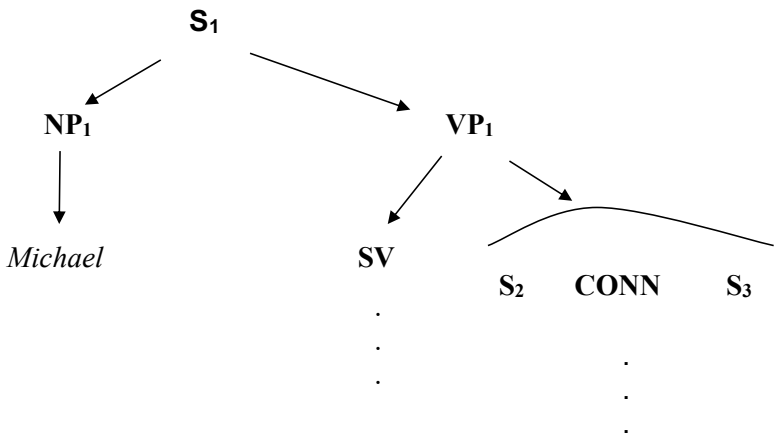
(2) Michael left the office. He had to attend to urgent business.

Prima facie, cross-sentential examples like (2) seem considerably different from intra-sentential examples like (1); once one ventures across multiple sentences, S_1 and S_2, nothing in S_1 can c-command anything in S_2. So, 'Michael' cannot c-command, and therefore bind, the pronoun 'he' in the succeeding sentence.

However, if one gives 'Michael' wide scope and rephrases the statement into a single sentence, one can bind the pronoun with its proper name antecedent. Here is another way to rephrase (2) without loss of meaning:

(2') Michael is such that he left the office and he had to attend to urgent business.

I will only give a partial tree diagram for (2'), but the result will sufficiently show that 'Michael' c-commands both instances of 'he':



The referent of both anaphoric instances of ‘he’ is given by the antecedent ‘Michael.’ However, it should be noted that the pronoun ‘he’ is not always an anaphor. For example, in the following case where ‘he’ is a demonstrative, one could be pointing to Oscar while telling Pam:

(3) Michael left the office. He [Oscar] had to attend to urgent business.

Nonetheless, the common intuition, independent of such demonstrative context, coheres with (2) and (2’); ‘he’ picks out the same referent as ‘Michael.’

Pronominal anaphora can sometimes present problems when the referent is not a specific entity but a generalized quantified expression. However, in many cases, one may resolve these issues through straightforward logical analysis. For example, take the following quantified statement:

(4) Every teacher reviews their syllabus.

The quantified antecedent ‘every teacher’ does not explicitly identify a singular referent for the pronoun ‘their’ with which it covaries. Nonetheless, a quantified translation of the statement shows that the pronoun, when expressed as a variable, is within the quantifier's scope and therefore bound by it:

(5) $\forall x[teacher(x) \rightarrow reviews(x, syllabus.of(x))]$

In this case, ‘their’ *covaries* with ‘every teacher’ since the content of the pronoun will change depending on which unique referent is assigned to ‘x’ in (5). Thus, within the scope of the given universal quantifier, the anaphoric expression readily translates as a bound variable. Specifically, for any person *x* such that *x* satisfies the condition of being a teacher, *that person x* reviews *x's* syllabus.

1.3: Troublesome Anaphora

However, instances of pronominal anaphora can sometimes present problems that a simple logical approach cannot solve. One such problematic case is *donkey anaphora*, a term coined by Peter Geach (1962). To be precise, donkey anaphora—also commonly known as “donkey

sentences"—refers to any instance in which a sentence contains an indefinite noun phrase within a relative clause (like an if-clause) and a pronoun outside that particular relative clause that is, nonetheless, an anaphor of that indefinite noun phrase (Heim 1982, p. 35).

Although Geach used examples of statements that normalize violence against donkeys, other examples of donkey anaphora can be used to demonstrate its inherent problem: donkey sentences complicate a linguistic understanding of how pronouns find their referents, as neither c-command nor standard binding principles fully account for the pronoun's reference, which covaries with the quantified antecedent. Thus, because on a standard logical analysis such pronouns are not bound, one cannot understand them as referring expressions (King and Lewis 2021, p. 9). For example, consider the following donkey sentence:

(6) If a gardener plants a tree, he waters it.[3]

A similar predicate logic translation for (6) yields the following result if literally translated:

(7) $\exists(x,y)[gardener(x) \wedge tree(y) \wedge plants(x,y)] \rightarrow waters(x,y)$

It is essential to note that while 'x' and 'y' are bound in the antecedent, the variables 'x' and 'y' contained in the consequent are not bound by the existential quantifier. If the antecedent satisfies some truth condition, the statement, as translated, cannot yield a truth value since the consequent cannot be readily assessed. So, although (7) is a well-formed formula, it is not a sentence due to the free-variable occurrences in its consequent.

Nonetheless, one might think it possible to modify the logical expression so that both variables 'x' and 'y' in the antecedent and consequent are bound by the existential quantifier. An attempt to do so would yield the following:

(8) $\exists(x,y)[(gardener(x) \wedge tree(y) \wedge plants(x,y)) \rightarrow waters(x,y)]$

Yet, this formalization must be revised, as it fails to represent the intended meaning accurately. To understand why this formulation is infelicitous, consider the following logical statement, which is equivalent to (8)[4]:

(9) $\exists(x,y)[\neg gardener(x) \vee \neg tree(y) \vee \neg plants(x,y) \vee waters(x,y)]$[5]

In other words, two objects exist, 'x' and 'y,' such that either 'x' is not a gardener or 'y' is not a tree, or 'x' does not plant 'y' or 'x' waters 'y.' (9) is true if and only if at least one of the disjuncts is true. Thus, an object within the domain of discourse that is, nonetheless, irrelevant to the context of an utterance of (6) could satisfy such a truth condition; a non-gardener guarantees the truth of (9). For example, an interpretation such that 'my computer' is assigned to 'x' maps (9)—and, by extension, (8)—to true. Such a weak interpretation, it would seem, should not be acceptable to capture the meaning of (6).

Other logical translations employing the existential quantifier fail similarly. Even if one were to restrict the existential quantifiers such that only gardeners and trees may be assigned to 'x' and 'y,' respectively, one would, it seems, still fail to capture the intended meaning. Consider the following logical statement:

(10) $\exists(x,y)[gardener(x) \wedge tree(y) \wedge (plants(x,y) \rightarrow waters(x,y))]$[6]

In standard English, (10) states there is a gardener and a tree and if that gardener plants that tree, then they water it. The problem with (10) is that it maps to true if there is a gardener and a tree and that gardener waters that tree but *does not necessarily plant it* (Beaver and Frazee, 2022, pp. 99-100). As with (9), consider an equivalent statement to (10):

(11) $\exists(x,y)[gardener(x) \wedge tree(y) \wedge (\neg plants(x,y) \vee waters(x,y))]$

As is evident and warrants no further explanation, *any* existential quantifier containing a conditional within its scope will fall victim to such infelicitous translations of (6).

Since an existential quantifier binding the variable does not appear to capture the intended meaning of such statements, one might think to formalize this example of donkey anaphora as a *universally quantified* statement as follows:

(12) $\forall(x,y)[(gardener(x) \wedge tree(y) \wedge plants(x,y)) \rightarrow waters(x,y)]$

For all gardeners who plant trees, those gardeners also water said trees. Certainly, employing ad hoc reasoning, one might argue that instances of donkey anaphora resemble *universally quantified statements in disguise*. Regardless of this ad hoc reasoning, the quintessential problem of donkey anaphora remains: the lack of a general principle for translating indefinites into either existentials or universals results in an unsatisfactory situation, as it prevents us from having a deterministic translation procedure (Beaver and Frazee 2022, p. 100).

SECTION 2: RUSSELLIAN INDEFINITE DESCRIPTIONS AND SEMANTIC REFERENCE

To illustrate the inevitable frustration stemming from the absence of a universal principle, I will examine Russell's theory of indefinite descriptions alongside Strawson's criticism of this approach. This analysis will reveal the challenges in developing a comprehensive framework encompassing referential and non-referential indefinite antecedents. Ultimately, if we understand anaphoric pronouns as variables that covary with their antecedents, then it follows that these pronouns do not necessarily have to refer to any real-world entity.

2.1: RUSSELL'S THEORY OF INDEFINITE DESCRIPTIONS

In virtue of A-form propositions, the set of gardeners who plant trees, as represented in (12), may be empty. So, if anaphoric pronouns pick out their referent from their antecedent phrase, and if the antecedent phrases of donkey sentences refer to non-existent members of an empty set, then which referent does the anaphoric pronoun pick out? If one accepts pronouns (in either the referential or attributive senses[7]) as picking out referents of non-empty sets, then proponents of a universal quantification treatment of donkey sentences must explain in a structured, algorithmic manner why anaphoric pronouns are treated differently from non-anaphoric ones.

Russell (1919) argued that indefinite descriptions fail to pick out any *particular* referent[8]; instead, indefinite noun phrases denote an arbitrary member of a non-empty set across all occasions of use. In the case of an ordinary subject-predicate statement involving an indefinite description, the intersection of the subject and predicate classes includes at least one member. To illustrate, consider the following sentence:

(13) A coyote wandered into my yard.

The intersection of the subject class—'coyotes'—and predicate class—'things that wandered into my yard'—is non-empty, but the expression does not refer to any particular member of the intersection of both classes. In this regard, Russell's treatment of indefinite noun phrases (or descriptions) contains existential import but lacks a *uniqueness* claim (i.e., it is not a genuinely referring expression); 'a coyote' denotes a member of the non-empty set 'coyotes' without picking out the referent of any particular coyote.

However, as Strawson (1950, pp. 341-342) demonstrated, Russell presupposes that such uses of indefinite noun phrases fail to refer across *all* contexts. Such a presupposition, Strawson argues, presents a false dilemma. As should be apparent from (13), the *same* indefinite noun phrase may succeed in picking out *different* unique referents depending on the given context; an utterance of (13) may be tokened on various occasions without necessarily referring to the same coyote. However, such token utterances classify 'a coyote' as a *genuinely referring expression*.

Strawson's objection to Russell's treatment of indefinite descriptions brings to light a semantic ambiguity of such expressions; some tokens of utterances with indefinite noun phrases may pick out a unique referent, as in (13). However, not all tokenings of such utterances necessitate reference. For example, consider these two anaphoric variants of (13):

(14) A coyote wandered into my yard. It is tearing through my trash and leaving a mess.

(15) A coyote wandered into my yard. It tore through my trash and left a mess.

In (14), 'a coyote' uniquely refers to a member of the class of coyotes, and, in virtue of 'it' being an anaphor, picks out its referent from the indefinite antecedent term. In other words, it is *that* coyote that wandered in my yard and *that* coyote that is leaving a garbage mess all over the yard. Conversely, it is not clear that 'a coyote' in (15) uniquely refers to any particular coyote; there may have been multiple coyotes that tore through the trash bags, or perhaps another animal or no animal is responsible for the garbage mess in the yard. For examples like (15), Strawson would presumably have to accept Russell's quantificational treatment of indefinite noun phrases. So, although Strawson's considerations are relevant to discussions of donkey anaphora, a formal semantic account of (15) is still necessary. Russell's theory of indefinite descriptions remains a viable candidate for instances where indefinite antecedents lack unique reference.

2.2: The Argument Against Anaphoric Pronominal Reference

So, it appears that anaphoric pronouns would be unable to pick out the same referent as an indefinite antecedent since such indefinite antecedents would fail to refer. Intuitively, it is clear that 'a coyote' in (15) serves as the antecedent for the anaphoric pronoun 'it,' but semantically, it is not evident how (or if) 'it' picks out a *unique* referent (or if it needs to).

The central argument identifying the problem could be formulated as follows. Call this, the *Argument Against Anaphoric Pronominal Reference*:

1. In cases like (15), indefinites serving as antecedents of anaphoric expressions do not have unique referents.
2. Anaphoric pronouns pick up their referents from their antecedents[9].
3. So, anaphoric pronouns pick up their reference from their antecedents and indefinites serving as antecedents of anaphoric expressions do not have unique referents.
4. If anaphoric pronouns pick up their reference from their antecedents and indefinites serving as antecedents of anaphoric expressions do not have unique referents, then anaphoric pronouns with indefinite antecedents do not have unique referents.
5. Therefore, anaphoric pronouns with indefinite antecedents do not have unique referents.

Intuitively, however, anaphoric pronouns *do* appear to have unique referents, or, at the very least, covary with their indefinite antecedent phrases. In other words, competent users of English—or any other language—seem to assume that uniqueness is built into the use of pronouns, whether deictic or anaphoric.

[1][10] was discussed in the previous section and, hopefully, was made clear that one has good reason to accept a Russellian treatment of indefinite antecedents in *some* anaphoric cases. After restricting the domain to instances of pronominal anaphora in which the indefinite antecedent does not appear to refer, one would have to refute [1] by explaining that these *prima facie* non-referential antecedents actually refer to a unique object. However, I will, for present purposes, accept [1] as true.

If one wishes to treat a statement like (15) from a purely semantic viewpoint, then one must reject either [2] or [4]; [3] is a logical consequence of [1] and [2] via conjunction introduction, and [5] from [3] and [4] via modus ponens. As will be expounded on in section III, Geach (1962) and other proponents of a bound-variables theory will reject [2]; Geach denies this necessary

connection between anaphora and picking up a reference in the world (Heim 1982, p. 18). Conversely, Evans (1977) will argue against Geach's treatment of donkey sentences and, by extension, deny [4][11]: anaphoric pronouns with indefinite antecedents *do,* in fact, have unique referents even if the conditions of [4] are met. This counterargument will be elaborated on further in section IV[12].

SECTION 3: BOUND VARIABLES

As shown previously, not all indefinite antecedents can be treated as referring expressions. So, as the Argument Against Anaphoric Pronominal Reference suggests, the anaphoric pronoun that covaries with such an indefinite antecedent itself cannot have a reference. For Geach (1962), there is nothing about anaphora that suggests its constituent parts must pick out a referent; one can easily reject the argument's conclusion by denying [2].

3.1: GEACH'S BOUND VARIABLES THEORY

Geach posits that both within a single sentence and across multiple sentences, anaphoric references can be directly converted into the language of first-order quantificational logic. He contends that the relationship between an anaphoric pronoun and its antecedent mirrors the relation of a variable to its quantifier in first-order predicate logic (p. 139). As detailed in subsection 1.2, untroublesome anaphoric expressions can be easily translated into the formal statements of predicate logic. Geach goes further, proposing that every case where the indefinite antecedent term fails to refer are expressible as quantified statements. To illustrate, let's revisit example (15):

(15) A coyote wandered into my yard. It tore through my trash and left a mess.

This cross-sentential anaphoric statement can be rephrased in the following way:

(15') A coyote is such that it wandered into my yard, tore through my trash, and left a mess.

(15) and (15') are equivalent and may be expressed in first-order predicate logic as follows:

(16) $\exists x[coyote(x) \wedge wand.yard(x) \wedge tore.trash(x) \wedge left.mess(x)]$

Geach's treatment of anaphoric pronouns as bound variables also appropriately translates indefinite quantified expressions using the determiner 'every.' For example, consider a showrunner for a Broadway play uttering the following anaphoric statement to their producers:

(17) Every actor in the play received high praise. They will be asked to work more often.

If one applies Geach's treatment of pronominal anaphora to (17), they will find that this statement does not lose meaning when rephrased as the following:

(18) Every actor in the play is such that they received high praise and will be asked to work more often.

Such a translation, therefore, is readily available for (18):

(19) $\forall x[play.actor(x) \rightarrow (praise(x) \wedge will.work.more(x))]$

Geach characterizes instances of anaphoric pronouns as "pronouns of laziness" where one uses pronominal expressions out of convenience rather than necessity; these pronouns stand in for indefinite or definite noun phrases that have been previously mentioned or can be readily inferred. These pronouns facilitate smooth discourse in a particular language without being essential to its meaning. Essentially, Geach suggests that using such anaphoric pronouns indicates a skilled language user's attempt to strike a balance between expressiveness and efficiency.

However, his stance on coreference becomes evident in the context of unbound pronominal anaphoric expressions involving quantified antecedents, such as donkey anaphora: "When [an anaphoric] pronoun is not a pronoun of laziness, it is in general quite absurd to treat it as a 'singular referring expression' and ask what it refers to" (p. 152). So, as stated in subsection 2.2, Geach denies that anaphoric pronouns—and, by extension, anaphora in general—must refer at all. More bluntly, Geach disregards the philosophers' and linguists' desires to explain such an intuitive understanding of anaphoric reference as "simply a prejudice or a blunder" (p. 153).

Thus, Geach's solution to the problem of donkey anaphora is simple: only singular antecedents like proper names and definite descriptions definitively pick out a referent to which

the corresponding anaphoric pronoun will refer. If one is presented with a donkey anaphoric statement where the antecedent is some quantified expression, then one must bite the bullet and admit that the anaphoric expression does not refer. As Geach states, indefinite antecedents of donkey anaphora contain "no criterion of identity, not at least in the widest sense… which alone is relevant to the… pronouns" (p. 170). Therefore, Geach successfully denies [2] from the Argument Against Anaphoric Pronominal Reference.

3.2: EVANS'S OBJECTION TO GEACH

In his influential article "Pronouns, Quantifiers, and Relative Clauses" (1977), Gareth Evans primarily argues against Geach's bound-variables treatment of pronominal anaphora. He proposes that certain instances of pronominal anaphora elude a simple quantificational translation as Geach suggests. Anaphoric pronouns, when covarying with indefinite quantifier antecedents, behave differently from bound variables. As such, Evans shows that certain bound-variable translations of anaphoric statements do not capture the original meaning.

Evans provides multiple examples in support of his claim that some anaphoric pronouns get their reference from indefinite antecedents that do not bind them, but I will only address one such example. Evans claims that "there are sentences containing pronouns, whose antecedents are quantifiers, but which are not naturally interpreted in the way that would result if the pronouns were bound by those quantifiers" (p. 109). In other words, there are sentences where, despite the presence of anaphoric pronouns with indefinite quantifiers as their antecedents, the natural interpretation of these sentences does not align with the pronouns being bound by those quantifiers. So, even though the structure of the sentence might imply that the pronoun should refer to something specific dictated by the quantifier, the way one actually understands the sentence does not follow Geach's treatment.

Consider the following two *seemingly* equivalent statements:

(20) A few relatives came to my party. They played Scrabble.

(21) A few relatives are such that they came to my party and played Scrabble.

Perhaps it is not immediately obvious why (20) and (21) denote semantically distinct statements. To see why this is the case, allow me to rephrase (20) using a definite quantifier:

(22) Exactly three relatives came to my party. They played Scrabble.

If Geach's treatment of anaphoric pronouns as bound variables of the indefinite antecedent is correct, then (22) can be rephrased as the following without loss of meaning:

(23) Exactly three relatives are such that they came to my party and played Scrabble.

However, (22) and (23) are not semantically equivalent. The intuitive reading of (22) states that *three and only three* relatives came to my party and *all three* of these relatives played Scrabble. The paraphrase represented in (23) does not necessitate this same meaning; it is possible that five relatives came to my party, but that only three of them played Scrabble (perhaps the other two thought their time would be better spent socializing). So, although, I used definite quantifiers to demonstrate this difference in truth conditions, the same distinction of truth conditions occurs for plural indefinite antecedents of donkey sentences.

Evans's refutation of Geach does not invalidate Geach's denial of [2]. However, Evans does demonstrate that a sufficient semantic theory accounting for donkey anaphora is not settled with a bound-variables treatment; a new theory must be presented with accounts for plural (but not maximal) indefinite antecedents.

SECTION 4: EVANS'S E-TYPE PRONOUNS

Unlike bound variables, Evans's treatment of e-type pronouns can be assigned a specific referent and evaluated independently for truth or falsehood within their immediate sentence contexts. In this sense, Evans claims, "such pronouns are not genuine singular terms in the sense in which ordinary proper names and demonstrative expressions are; rather they are singular terms whose reference is fixed by description" (p. 104). As Evans stipulates, e-type anaphoric pronouns are those pronominal expressions which pick out the same referent as the indefinite antecedent by fixing an associated *definite* description to the pronoun. So, consider again (20):

(20) A few relatives came to my party. They played Scrabble.

According to Evans, (20) may be rephrased as the following:

(24) A few relatives came to my party. [The few relatives that came to my party] played Scrabble.

The substituted definite description that uniquely picks out the referents of the indefinite antecedent noun phrase denotes the role of the anaphoric pronoun.

After addressing Geach's bound-variables treatment[13], Evans puts forward his theory of e-type pronouns: these "pieces of evidence seem to me to point in the direction of treating these pronouns… as singular terms whose denotation is fixed by a description recoverable from the clause containing the quantifier antecedent" (p. 111). However, Evans does not equate this formulation of an e-type pronoun to some particular definite description that picks out some unique referent. According to Evans, treating such pronominal anaphoric expressions as mere disguised definite descriptions results in ambiguity (p. 132).

Instead, Evans adopts Saul Kripke's reference-fixing of a singular term via some definite description to explain the role of E-type pronouns in pronominal anaphora. Kripke (1981) uses 'one meter' as an example of an object which is rigidly designated through the use of a particular description:

> There is no conflict between the counterfactual statement and the definition of 'one meter' as 'the length of *S* at t_0' [...] [This] does *not* say that the phrase 'one meter' is to be *synonymous…* with the phrase 'the length of *S* at t_0,' but rather that we have *determined the reference* of the phrase 'one meter' by stipulating that 'one meter' is to be a *rigid* designator of the length which is in fact the length of *S* at t_0. (p. 56)

In the case of "one meter" or some other rigid term whose reference is fixed via description, the referent of the term is contingent—the same stick *S* or empirical evidence used to define 'one meter' may have had a different length at time t_0—and known *a priori*—because we stipulate that 'one meter' refers to the length of *S* at t_0, we can know it independently of experience. For Evans,

fixing the referent of an anaphoric pronoun via definite description admits of this contingent *a priori* metaphysical and epistemological status regarding pronominal use.

Evans's adaptation of Kripke's reference fixing via description renders his position on the role of e-type pronouns, and his denial of [4], clear: "I wish to emphasize that *one is in no way committed, by saying that e-type pronouns have a referential role, to the view that their quantifier antecedents refer*." In fact, if there does exist a member of the particular quantified set in question, then the "pronouns denote the items … that *verify* the quantified sentence" (p. 124). In this sense, the anaphoric pronouns in all donkey sentences are accounted for; the pronoun may be thought of as some contingent *a priori* statement whose referent is fixed via definite description[14].

The anaphoric pronoun refers to the quantified antecedent expression, but this fact does not imply that the indefinite quantified antecedent is genuinely referential: "Precisely because the term has its reference fixed by description, its reference may be specified, and therefore the truth conditions of any sentence containing it may be specified, whether or not it has a reference" (p. 124). Effectively, Evans provides a *static* semantic explanation for our intuitive, natural understanding of donkey anaphora and other pronominal expressions.

SECTION 5: PROBLEMS WITH EVANS'S ACCOUNT

Evans assumes that his adapting Kripke's reference-fixing strategy allows e-type pronouns to fix its referent to its antecedent, even if the antecedent contains an indefinite quantified expression. However, nothing in Kripke's thoughts on referent-fixing via description nor contingent statements knowable *a priori* appear to demonstrate this connection. Certainly, if a description *does* pick out the same reference as its antecedent, then Evans's theory of e-type pronouns will prove to be an adequate semantic account. Nonetheless, if anaphoric pronouns do act as variables and those variables are unbound within a quantified statement, what semantic rule allows a particular definite description to be substituted for such a variable without running into the same problem of reference as unbound pronouns?

For example, refer back to (6) from the first section of this paper:

(6) If a gardener plants a tree, he waters it.

Using an e-type analysis, two definite descriptions—one for 'he' and one for 'it'—may be substituted into the expression:

(25) If a gardener plants a tree, [the gardener who planted the tree] waters [the same tree that the same gardener planted].

However, notice that if *two* gardeners plant *two* separate trees, then the same anaphoric statement would apply to both. However, from a purely semantic viewpoint, how can one determine to which gardener the utterer of the anaphoric expression refers? It appears there is not a convincing answer independent of contextual clues. So, although Kripke's referent-fixing via description may work for unbound anaphoric expressions with *one unique* referent, the tactic appears to fall short when quantified antecedents contain more than one member in their extension. Evans's assumption of fixing the reference via definite description to a quantified antecedent leads to semantic underdetermination; some pragmatic or psychological account of anaphora is also required to account for this underdetermination.

As Evans saw his theory, e-type pronouns had the added benefit of not appealing to mental events such as speaker intention. While Evans acknowledges that the context of utterance is relevant for fixing the referent via an apt definite description, the semantics alone of e-type pronouns are able to provide explanation for pronominal anaphora without appeal to pragmatics or "psychologizing": "And so we might find ourselves trying to determine the truth value… by asking who [or what] a man would have in mind when he uttered or wrote down the sentence. [...] On the view of E-type pronouns I have so far outlined there is… no license to engage in 'psychologizing'" (p. 129). However, recent developments in philosophical and linguistic literature suggest that psychological experience is necessary for the study of semantics (Pitt 2018) and that mental representations provide the best framework to understand general discourse anaphora[15]. Discussing the necessity of mental representation in any theory of anaphora is a valuable project in its own right, but one which lies beyond the scope of this paper. However, a preliminary dive into the topic appears to show evidence against Evans's descriptivist account and others influenced by it in favor of those that embrace mental representation in their theories of anaphora.

CONCLUSION

I attempted to thoroughly analyze the nuances of donkey anaphora and the competing theories of Peter Geach and Gareth Evans. Geach's bound variables theory posits that anaphoric pronouns act as bound variables that covary with their quantified antecedents. As a result, anaphoric expressions with indefinite antecedents fail to refer, a view that aligns with traditional logical analysis but struggles with the complexities of donkey anaphora. In contrast, Evans's e-type pronouns theory, influenced by Saul Kripke's ideas on referent-fixing and contingent a priori knowledge, argues that these pronouns, though associated with quantifier antecedents, are not bound variables. Instead, they refer through definite descriptive means, even if the indefinite antecedent fails to refer. This treatment offers a different perspective on the semantic underpinnings of pronominal anaphora.

However, Evans's theory, with its innovative application of descriptivism, addresses some of these shortcomings but still leaves questions unanswered, particularly regarding the semantic underdetermination and the role of context in fixing reference. The exploration of these theories underscores the complexity of linguistic phenomena and the ongoing challenge in linguistics and philosophy of language to develop comprehensive theories that adequately explain the interplay of syntax, semantics, and logic. This analysis invites further inquiry into the nature of anaphora, encouraging analysis of other theories of pronominal anaphora, such as discourse representation theory and other theories of dynamic semantics, which incorporate mental representations into their models.

NOTES

1. Henceforth, I will abbreviate terms denoting nodes for brevity.
2. A tree diagram would work as well, but it is much simpler and consumes less space to get the point across effectively.
3. I use an example of a *conditional* donkey sentence, but one could also use a *relative clause* donkey sentence as follows and still arrive at the same problematic result: "Every gardener who plants a tree waters it." The issue, in this case, arises with the pronoun 'it' picking out the same referent as the indefinite noun phrase 'a tree,' which does not denote any particular tree. Additionally, we could phrase (6) as two simple sentences, but the problem still arises: "A gardener plants a tree. He waters it."

4. For the present discussion, I am assuming the material conditional to be true.
5. Assume distinctness of objects 'x' and 'y' (i.e., 'x' is not equivalent to 'y').
6. As in (9), assume distinctness of objects.
7. See Donnellan (1966). Although Donnellan objects to both Russell's and Strawson's use of *definite descriptions*, his treatment of such descriptions referential and attributive uses also seems to apply to uses of pronouns and indefinite descriptions as well. Kaplan (1977) expands on the use of indexicals in general, claiming that indexicals not only pick out a particular referent, but also act as *rigid designators*; pronouns and other indexicals pick out the same unique referent in *all possible worlds.*
8. See Ch. 16, in particular Russell's explanation of why 'a man' in 'I met a man' fails to refer.
9. In other words, anaphoric pronouns are *referring expressions* in that their reference depends on the antecedent expression. This premise may also be construed as saying that anaphora, in general, picks out reference, as Geach argues against.
10. To refer to the premises in this argument, I will use the premise number enclosed within brackets. So, for example, I will refer to the first premise as '[1],' the second premise as '[2],' and so on.
11. See Neale (1990), especially Ch. 5.
12. Of course, as I will argue in section V, Evans's theory of e-type pronouns fails to establish anaphoric pronominal reference.
13. Regarding the Argument Against Anaphoric Pronominal Reference, Evans believes [4] to be the faulty premise. However, as I argued earlier, I think Geach successfully refutes [2], and Evans's critique of Geach is regarding his *treatment* of donkey anaphora, not Geach's argument against [2].
14. Evans expounds further on these points in "Pronouns" (1980) and "Reference and Contingency" (1979).
15. See general discussions on "dynamic semantics" and "discourse representation theory" in Geurts et al. (2020)

BIBLIOGRAPHY

Beaver, David, and Joey Frazee. (2022) "Semantics," in: R. Mitkov (ed.), *The Oxford Handbook of Computational Linguistics,* 2nd edition (Oxford: Oxford University Press)

Donnellan, Keith. (1966) “Reference and Definite Descriptions,” *The Philosophical Review* 75(3), pp. 281-304

Evans, Gareth. (1985) “Pronouns,” in: A. Phillips (ed.), *Collected Papers,* pp. 214-248 (Oxford: Oxford University Press)

Evans, Gareth. (1985) “Pronouns, Quantifiers, and Relative Clauses,” in: A. Phillips (ed.), *Collected Papers,* pp. 76-152 (Oxford: Oxford University Press)

Evans, Gareth. (1985) “Reference and Contingency,” in: A. Phillips (ed.), *Collected Papers,* pp. 178-213 (Oxford: Oxford University Press)

Geach, Peter. (1980) *Reference and Generality,* 3rd edition (Ithaca, NY: Cornell University Press).

Geurts, Bart, David I. Beaver, and Emar Maier. (2023) “Discourse Representation Theory,” in: E. N. Zalta and U. Nodelman (eds.), *Stanford Encyclopedia of Philosophy* Winter 2023 , https://plato.stanford.edu/archives/win2023/entries/discourse-representation-theory/

Heim, Irene. (1982) “The Semantic of Definite and Indefinite Noun Phrases,” Dissertation

King, Jeffrey C., and Karen S. Lewis. (2021) “Anaphora,” *Stanford Encyclopedia of Philosophy* Winter 2021, https://plato.stanford.edu/archives/win2021/entries/anaphora/

Kripke, Saul. (1981) *Naming and Necessity* (Malden, MA: Blackwell Publishing)

Neale, Stephen. (1990) “Pronouns and Anaphora,” *Descriptions* (Cambridge, MA: MIT Press)

Pitt, David. (2018) “What Kind of Science is Linguistics?” in: C. Behme and M. Neef (eds.), *Essays on Linguistic Realism*, pp. 7-20 (Amsterdam: John Benjamins Publishing Company)

Russell, Bertrand. (1919) *Introduction to Mathematical Philosophy* (London: Allen and Unwin)

Strawson, Peter. (1950) “On Referring,” *Mind* 59, pp. 320-344

ADVERBIALISM AND THE MANY-PROPERTY PROBLEM

Ziang Li

INTRODUCTION

Adverbialism was originally proposed by R. M. Chisholm (Chisholm 1957, pp. 142-167), C. J. Ducasse (Ducasse 1968, pp. 192-217), and W. Sellars (Sellars 1968, pp. 167-169).[1] Sellars proposed that to have a sense impression is to be "impressed in a certain manner" (Sellars 1968, p. 167). For instance, consider the following sentence:

(1) S has a sensation of a red triangle.

(1) can be paraphrased as (2):

(2) S has an of-a-red-triangular sensation.

(2) can be paraphrased as (3):

(3) S senses in an of-a-red-triangular manner.

As (3) suggests, adverbialism postulates perception as a certain mode or way that the subject perceives. The main motivation for adverbialism is to develop a theory that can avoid committing to non-physical entities such as sense data, in place of sense-datum theory. According to sense-datum theory, the perceptual process has an act-object structure, with sensation being the action and sense data being the object, and perception is the relationship between sensation and sense data. However, sense-datum theory commits to the existence of sense data as non-physical entities, which costs ontological economy and raises many problems. Adverbialism can avoid committing to non-physical entities between subjects and objects. It only commits to the existence of events of subjects perceiving which can be modified adverbially. For example,[2] on the face of it, "Jones has an intense pain" has the same logical structure as "Jones has a ferocious dog," which leads people to think that "pain" is the object, just like "a ferocious dog". However, this is actually a misunderstanding caused by our

language expression. We can eliminate this misunderstanding by reformulating the sentence. "Jones has an intense pain" can be modified to "Jones is pained intensely." By modifying "an intense pain" from the predicate to an adverb modifying "Jones's sensation," "pain" is no longer the object of the subjective experience of pain, but is sensed in a painful way. This means that a certain subjective experience is not a relation between sensation and a specific non-physical object, but subject experiencing in a specific way. In other words, when we talk about what it feels like, we are not referring to sensing a certain non-physical object, but rather describing what the current psychological state or ongoing mental process of the subject having that sensation is like. Therefore, it is tendentious to interpret adverbial theory merely as a linguistic claim or translation maneuver, because its very essence is a metaphysical statement about the nature of perceptual experience (Crane & French 2021).[3]

However, Frank Jackson (1975) issues a challenge for adverbialism, which he calls the "Many-Property Problem" (henceforth abbreviated as the "Problem"). The purported problem is that adverbialism will fail to distinguish between perceptual experiences in which diverse properties are instantiated jointly. It reveals that the intuitive paraphrasing of adverbialism invites indeterminacy of the adverbs because there can be structurally equivalent adverbial paraphrasing for distinct perceptual experience. Current studies in the philosophy of perception typically exclude adverbialism because of its seemingly unpromising explanatory power in addressing the Problem. However, several philosophers have recently revaluated and revived adverbialism from the longstanding predicament (e.g., Kriegel 2008; Banick 2018; D'Ambrosio 2019, 2021; Gert 2021). Another notable development of this issue is that Crane and Grzankowski (2022) claim that the Problem may arise for other widely accepted theories of perceptual experience, especially intentionalism and naïve realism. When confronted with the Problem, these theories must appeal to some *ad hoc* explanations to account for the properties that are jointly represented in perception.

In section 1, I illustrate the Problem and its variant for adverbialism. I will discuss two solutions from Tye (1975, 1984a) and D'Ambrosio (2019, 2021).[4] Then, in section 2, I will discuss whether the Problem poses a challenge to other current theories in the philosophy of perception. In section 3, I will propose a causal theory to tackle the preceding problems while indicating that this approach invites three general troubles, which are concerned with its

metaphysical foundation and explanatory power.

SECTION 1: PROBLEMS FOR ADVERBIALISM AND SOLUTIONS

In this section, I will demonstrate the Many-property problems and its variant, the Many-relations Problem. Although Tye (1975, 1984a) and D'Ambrosio (2019, 2021) can resolve the Problem respectively, their adverbial theories may have potential limitations concerning specific perceptual experience.

1.1 MANY-PROPERTY PROBLEM

The Many-Property Problem includes two parts: the expressive problem and the inferential problem.[5] According to the Problem[6], it is problematic for adverbialists to translate perceptual reports as follows:

(4) S hallucinates a red triangle and a green square.[7]

Based on the basic intuition from adverbialism, (4) can be paraphrased as any of the following three statements:

(5) S hallucinates redly and triangularly and greenly and squarely.
(6) S hallucinates redly triangularly and greenly squarely.
(7) S hallucinates red-triangular-ly and green-square-ly.

Moreover, (8) S hallucinates a green triangle and a red square.

The expressive problem is that if we follow the paraphrasing of (5), then (8) can be paraphrased as:

(9) S hallucinates greenly and triangularly and redly and squarely.

(5) and (9) are logically identical. However, (4) is not identical with (8). Hence, adverbial paraphrasing fails. What if we follow the structure of (6)? Adopters of (6) must provide principled reasons to distinguish which adverbs modify the other one. Otherwise, it

would be arbitrary to decide the priority of adverbs in explanation.

If we take (7) as the correct form of adverbialism, it will lead to the inferential problem. Because (7) cannot entail:

(10) S hallucinates redly.

Since (7) uses "fused" modifiers, the inference from (7) to (10) would not preserve the compositionality in its original sense because it is a simple, non-derivative expression. For example:

(11) Turtles move excessively-slow-ly.

Clearly, (11) cannot entail:

(12) Turtles move excessively.

As D'Ambrosio (2019) points out, these problems indicate that adverbialism "lacks the expressive resources to distinguish between perceptual situations, or it lacks the logical resources to validate intuitive inference patterns" (p. 419).

1.2 TYE'S SOLUTION

To begin with, Tye (1975) and Sellars (1975) points out that Jackson neglects the "of-an-X-Y manner" paraphrasing. Because Jackson only discusses that the adverbial analyses of (5), (6), and (7) incur the Many-Property Problem, whereas he omits the form in (3). Based on the form of "of-an-X-Y manner", they argue that the Problem threatens nothing because the expressive problem will automatically resolve and the inferential problem can be easily tackled by logical regimentation.[8]

Furthermore, Tye (1984a) provides a more detailed version of adverbialism.[9] He introduces a two-place operator called "Coin,"[10] which operates in mapping F-ly and G-ly onto the compound function F-coincidental with G-ly. Thus, (4) is paraphrased as RCoinT(Hs) & GCoinSq(Hs).[11] In combination with several logic rules, he claims that the Problem has been

solved. Because through this operator, the structure of (4) and (8) is no longer identical, and it can generate a complex, derivative mode of sensing F-ly and G-ly, and then the inference from FCoinG(Hs) to F(Hs) is valid. However, I find an explicit definition of "Coin" is missing; it is not clear whether "Coin" is a purely logical operator, "a logically useful *façon de parler*" (Sellars 1975, p. 156), or it commits to a dubious metaphysical relation to hang coincident properties. There is no significant structural difference between "Coin" and the hyphenation within the intuitive paraphrasing. Tye needs further argument to support its function in distinguishing between different perceptual reports. Furthermore, as mentioned above, Tye's approach is challenged by the Many-Relations Problem, which I will discuss in the following section.

1.3 MANY-RELATIONS PROBLEM

The Many-Relations Problem is a strengthened variant of the Many-Property Problem which primarily attacks Tye's version of adverbialism. Consider the following cases (Dinges 2015, p. 235):

(13) Jane has a red and a green afterimage, and the former is brighter than and to the left of the latter.

(14) Jane has a red afterimage that is brighter than a green one and to the left of a yellow one.

According to Tye's regimentation, (13) can be paraphrased as:

(13a) R*Bri*GAj and R*Left*GAj,

where "*Bri*" and "*Left*" represent the operators like "Coin." However, (13a) is inadequate, since it leaves room for another red afterimage which is to the left of the green one but not brighter than it. Adverbialists can introduce a higher-level operator to map higher-level properties onto the compound function; this allows them to profess that the genuine form would be:

(13b) R(BriCoinLeft)GAj.

Regrettably, this move will still be inadequate for a situation like (14), because the third afterimage remains: “This operator helps only if there is a single pair of afterimages instantiating more than one relation” (Dinges *op. cit.*, p. 236), Tye’s “Coin”, thereby, loses its value[12].

1.4 D’AMBROSIO’S SOLUTION

D’Ambrosio (2019, 2021) proposes a new regimentation to solve the two problems above. He offers two methods to develop the semantics compositionally: the type-shift and semantic incorporation.[13] The former refers to the quantifier phrase or general quantifier bringing its value in general semantics into its restricted domain, i.e., into a certain property. The latter is a description of a semantic phenomenon; when semantic incorporation occurs, semantically independent nouns are incorporated into verbs, forming a compound transitive verb, in which the noun serves as a modifier to restrict the extension of the verb. D’Ambrosio attempts to establish an adverbial semantics, and prove upward monotonicity[14] in virtue of the semantics. By semantic incorporation, (4) can be reconstructed into:

(4a) ∃*e*(hallucinating(*e*) & agent(*e*, S) & Char (*e*, [[a red triangular patch and a green square patch]])) (D’Ambrosio 2019. p. 433).

Here, “*e*” represents a perceptual event, “S” represents the subject, and “Char” is short for “characterization”, which is “posited specifically to capture the semantic contribution of quantificational modifiers” (*ibid.*, p. 424). More specifically, “an event of perceiving e is characterized by the quantifier [[*x*]] iff for every event *e* ¯that makes *e* successful, *e* ¯has some subevent *e′* such that *e′* is a relational perception, and an *x* is such that *e′* has *x* as a theme” (*ibid.*, p. 429).

In terms of type-shift, (4) can be reconstructed into:

(4b) (a red triangle and a green square)-ly (hallucinate) (S) (*w*) (*t*) = iff there is an act of hallucinating e in w at t of which S is the agent such that the property of being seen

by S is an element of $\lambda P[\exists x \exists y$[red(x) $\wedge$ triangular(x) $\wedge$ green(y) $\wedge$ square(y) $\wedge$ P(x) $\wedge$ P(y)] at (w', t'), for any worlds w' and times t' such that (w, t) T (w', t') (D'Ambrosio 2021, p. 827).

Or, in less technical vocabulary: When S hallucinates (a red triangle and a green square)-ly, given the analysis, every world in which S's hallucinating is successful is one in which S hallucinates a red triangle and a green square. But every world in which S hallucinates a red triangle and a green square is one in which S hallucinates a square. Thus, every world in which her hallucinating is successful is one in which S hallucinates a square.

The general idea of (4a) and (4b) is that for any event e concerning perception, of which x is the agent, it is characterized by "Char (e, [[ ]]))," the perceived objects. D'Ambrosio claims that both approaches can satisfactorily resolve the Many-Property Problem. Evidently, after his paraphrasing, (4) is different from (8) (after the same operation) and validates the inference form (4) to (10) because of the relation between events.

I have two reservations about his approach.[15] First, some more particular and subjective experiences may be difficult to characterize by "Char," such as blurriness and sharpness. For example, when two people both see a (blur-red-triangle)-ly, the degree of blurriness in their modes of seeing may be different. Capturing such differences through "Char" may pose the following two problems: first, it is unclear to which sub-event the blurriness is attributed to; second, even if there is a clear attribution, it may be difficult to map the differences in the degree of blurriness into "Char" as a generalized quantifier.

Secondly, veridical hallucination[16] has been ruled out in D'Ambrosio's schema. He endorses the view that takes "perception to be explanatorily primitive and takes perception to be an activity whose aim is to perceive relationally." In addition, "successful and unsuccessful perceptual events have a common factor, and that this factor is a person-level perceptual event" (D'Ambrosio 2019, 51 f.). Although he appeals to a success condition to avoid being reduced to relationalism, the veridical hallucination, in effect, satisfies the success condition since it coincides with the condition that leads to veridical experience. Therefore, how is it possible that his "characterization" characterizes one's visual experience without appealing to a relational view, given that what is happening at the first-person-level is inconsistent with the

actual situation?

SECTION 2: THE MANY-PROPERTY PROBLEM AND PERCEPTUAL EXPERIENCE

In general, the Many-Property Problem and its variants present dilemmas for adverbialism exclusively. Intriguingly, Crane and Grzankowski (2021) claim that when confronted with this problem, intentionalism and naïve realism must appeal to some *ad hoc* explanations to account for the properties that are jointly represented in perception, revealing their neglect of the dependence of perceptual experience on the Laws of Appearance (hereafter abbreviated as the Laws). The Laws refer to a term coined by Adam Pautz to describe a set of general principles that govern visual experience, and the existence of these laws can provide a metaphysical basis for the possibility of visual experience. They are "metaphysically necessary prohibitions on how things might appear" (Pautz 2017, p. 37).

Crane and Grzankowski argue that the Problem also arises for other theories of perceptual experience. For intentionalism, *x* is experiencing a red triangle and a green square such that *x* is visually representing that there is something red and triangular and something green and square. Intentionalism faces the inferential problem from the Problem because contents that are irrelevant to one's perceptual experience can be packed into the entailment from the contents of one's representational state: "the content of one's representational state entails some other content, it does not follow that one is thereby in a representational state with the entailed content" (Crane & Grzankowski 2022, p. 173). Therefore, a brute fact in intentionalism has not been explained away; that is, they allow specific entailments between the specific representational contents of the relevant perceptual states to explain some connections between experiences, while blocking other connections which also can be entailed from the representational contents.

Furthermore, for the naïve realists, perceptual experiences are determined by external factors that are independent of the perceiver; *x* is experiencing a red triangle and a green square such that there are a red triangular object and a green square object that are perceived by *x*. Crane and Grzankowski argue that there are many properties of perceptual objects that cannot be perceived through common experience, such as the molecular properties of objects, but these properties also belong to the perceived objects. Hence, naïve realism needs to explain why

specific properties of perception and its objects are connected, and not others. They attribute the shared problems between intentionalism and naïve realism to the Laws: "why some contents must hang together and others do not" (*ibid*, p. 174).

While Crane and Grzankowski pose a challenge to the mainstream theory of perceptual experience to some extent, their arguments are neither sufficient nor complete. At least two loopholes exist. First, they lack explanations of representational content. There are two basic positions regarding the nature of representational content: internalism and externalism. The former suggests that the content of perceptual experience is determined by the internal physical states of the experiencing subject. Conversely, the latter suggests that the content of perceptual experience is at least partly determined by external relations, and that two subjects with the same internal physical state may have different perceptual contents. These two positions partially respond to the inferential problem relating to intentionalism; regardless of the position taken for representational content, the representational content experienced by the subject can be fixed by internal or external physical causations. However, whether appealing to causation can satisfactorily explain away the laws of appearance requires further discussion.

Second, they overlook the naïve realist's conception of "acquaintance," a relation between subjects and external objects. Naïve realism holds that the phenomenal properties of perception are fixed by the subject being directly acquainted with mind-independent physical objects, and that the ability of the perceptual subject to be directly acquainted with physical objects is a fundamental feature of perceptual relations. Thus, proponents of naïve realism can argue that certain properties of the experienced object are excluded from perceptual experience because the subject cannot have the direct acquaintance with these properties in ordinary experience, and therefore, these properties are ruled out in our perceptual experiences. However, naïve realists still require a more in-depth analysis of the concept of acquaintance, otherwise it will end up being another brute fact about appearance. If acquaintance is considered an irreducible relation, then this leaves unexplained facts for naïve realism.

SECTION 3: THE CAUSAL THEORY OF ADVERBIALISM

Before digging into the issue, I need to briefly clarify two crucial points. It seems bizarre to attempts to develop a causal theory for adverbialism. According to the traditional view,

adverbial analysis eliminates the experiential relationship between subject-object and treats perceptual states as non-relational states. However, I argue that this is not a fixed orthodoxy for adverbialism because I take cause-and-effect not as a relation in the metaphysics of perception but as an explanatory mode in my adverbial theory. Consequently, it is a causal theory in a *quasi*-adverbialism manner. Within it, cause-and-effect modifies the way that subjects perceive just as adverbs do, but at a metaphysical level. It contains facts about the presence of our perceptual experiences. The prefix *quasi* suggests that I have no intention of eliminating my argument from ontological commitment to causation. This paper merely utilizes the notion of causation to illustrate adverbialism better. Further, it might fail to generalize this theory to the causal theory of perception, because the core of this thesis is the exploration of the causal theory of adverbialism as opposed to the causal theory of perception. What I intend to provide is a schema to wipe out the confusing explanatory mode behind adverbialism in virtue of causation. I have no intention to invoke the mysterious and knotty metaphysical problems that causation might incur, notwithstanding the fundamental role it plays. In addition, I will not endorse a causal form of direct realism as what Crane and French suggested adverbialists embrace (cf. Crane & French 2021). The traditional causal theory of perception treats veridical, illusory, and hallucinatory experiences as causally distinct cases, whereas the *quasi*-adverbial causal theory considers these cases as equally causally explicable (cf. Grice 1961; Tye 1982). In line with adverbialism, veridical perception and non-veridical perception share the same mental nature. Thus, why can they not be equal in the causal explanation? To be precise, equal in the causal explanation means that different causes can satisfactorily explain why subjects perceive things in such and such ways.

To begin with, I employ the causal analysis of seeing made by Tye (1982). For adverbialism, I align with D'Ambrosio (2019, 2021), adopting a view introducing characterization into adverbial paraphrase. Here are my definitions of the causal theory of adverbialism:

P senses x = df (i) There is a causal chain of events e which ends with P's having an experience of a sensation/stimulus S; (ii) within e some events involving P's sense organs, e.g., eyes and ears, causally *intervene between* some event involving x and P's experiencing S; (iii) within possible non-standard events, e.g., the veridical hallucination, other factors, e.g.,

ambient light and factitious appearances, causally *contribute to* some events involving *x* and P's experiencing S. (iv) *e* is characterized by [C][17] iff that event is causally explicable that the presence of *x* is due to [C].

Before I carry on, a few clarifications are necessary. "Intervene between" means that sense organs are fundamental to sensations while "contribute to" denotes that other factors are constitutive to non-standard sensations; both are indispensable for causing respective experience. Hence, sense organs and other factors are equal in causal explanation. "Causally explicable" means that in a charactered event, the presence of *x* is explained by abductive analysis and [C] contains causal information about the presence of *x*. An event of perception is characterized by [C] iff that event contains the causal chain that the presence of C is highly likely. My definition of [C] is different from D'Ambrosio's definition of "Char" because success conditions (the presence of the schematic objects) are embedded in "Char," while [C] takes causal information as explanatory basics of the presence of the schematic objects.

Now, that is enough for the preliminary. One may wonder how such analysis finesses the thorny issues raised above. As for the Problem, (4) and (8) are utterly heterogeneous cases because it is very likely that two distinctive notions of [C] characterize different hallucinations, and they are causally explicable in different ways. The second horn of this dilemma seems a tricky one, for [C] is different in (4) and (10); actually, what characterizes (4) is very likely sufficient for characterizing (10) because the causal chain leading to (4) could very likely lead to (10) at some point. Thus, (4) can entail (10). Concerning the many-relations problem, [C] has ruled out the possibility of the third sensum in sensation,[18] because [C] will be totally different if one conspires to introduce the third sensum.

The explanatory generalization can be reached regarding the veridical hallucination and the phantom limb pain case. Due to the crucial [C], the tension between [C] and present x is eased. For subjects can learn information from [C] without relinquishing their original experience.

Above all of these, the causal theory of adverbialism has at least two advantages. First, it resolves previous problems successfully and shuns dubious entities or relations, which is in line with the adverbialist's initial motivation. Second, it resorts to fewer technical apparatus than others, by only appealing to causal explanation. And compared with the structural

approach, the perceptual reports characterized by [C] are more accessible to intuition. Tye and D'Ambrosio's strategies emphasize the validity of formal semantics in adverbial paraphrasing, and overlook the actual perceptual experiential process. They deviate from the essence of adverbial theory, as adverbial theory emphasizes the specific modes of perceiving, while their theory shifts the focus to the ways in which adverbs modify perceptual activities.

Nevertheless, at least three possible responses need to be considered. Others may argue that, as Banick (2018) commented on Kriegel (2008), "it would be entirely *ad hoc* to suppose we could cherry-pick the one principle without getting the other" (p. 678). Unlike the determinable-determinate relation, the cause-effect relation is much more fundamental and ubiquitous. I appeal to causation mainly because it concerns facts about how we perceive the world around us.

Another response is a general one, and is concerned with the causal explanation. Admittedly, it is a substantial question to what extent empirical causation can explain the perceptual experience. The causal relation, or the fact that an event is causally reducible, implies that certain external factors become linked to the entire perceptual event. The molecular structure of an object is causally related to the production of our perceptual experience, but to what extent can the molecular structure influence the perceptual experience itself? Nonetheless, addressing this issue is beyond the scope of this paper.

Lastly, I added to D'Ambrosio's theory a causal chain that makes a sensation uniquely correspond to the nature of an object but there is a problem: each sensation has a causal chain of its own, regardless of whether the sensation is a real perception, an illusion, or a hallucination. The chain of causation does not appear to be as meaningful in distinguishing real experience from false or illusory experience. In other words, we could say that I perceived a banana yellowly and curvedly, and adverbs like yellowly and curvedly would not be used to describe the experience of an orange. However, according to the causal interpretation of adverbialism, I perceive a banana causally-yellowly-curvedly. Similarly, I sense an orange causally-orangely-curvedly, and the causal connection does not help me to distinguish the experience of a banana from the experience of an orange at the personal level. I would argue that the causal chain can be used to individualize personal experience, because it is constituted by many factors from the environment and brain states, which are helpful to distinguishing perceptual experience. If

such factors are identical, the two perceptual experiences would be indistinguishable. However, I admit that the causal theory of adverbialism cannot be utilized to capture the phenomenal characters of our experiences, and it requires further improvement.

CONCLUSION

Adverbialism has been reported out-of-date due to the dreadful Many-Property Problem and its enhanced version. Nevertheless, current literature has shown this view to be an improper stereotype. Conversely, "Adverbialism in philosophy of perception may be on the verge of a comeback" (Gert 2021, p. 699). Despite the fact that its current revitalization is not as impressive as it claims, it still informs us of the various possibilities that adverbialism may hold. The previous adverbial theory has overly emphasized the semantic-logical validity of perception, pre-setting all the details possessed by the perceptual objects, and only then inquiring about how they are presented, thus overlooking the rich process through which the world gradually unfolds in perception. This paper elucidated a new mode of adverbialism and assessed it within various dilemmas. This approach introduces a causal explanation into formal semantics characteristic of perceptual experience. However, it introduces two equally challenging issues: the issue of the efficacy of causal explanation and the status of causal relations in perception. Further issues remain to be expounded on.

NOTES

1. For a detailed reading of Chisholm and Sellars, see Casullo 1983, pp. 144-158. And the interpretation of Ducasse can be found in Butchvarov 1980, pp. 266-269. Additionally, a recent understanding of Sellarsian claims can be found in Corti 2022, which draws the convergence between Sellarsian and contemporary adverbialists (Corti 2022, pp. 14-15).
2. The examples are from Tye 1984b, p. 319.
3. However, Steenburgh argues that adverbial statement "translates trivially, or worse, abbreviates" (Steenburgh 1987, p. 380).
4. Kriegel (2008) applies the determinable-determinate relation to tackle the Problem. However, Grzankowski (2018) objected this line of thought, claiming that that "this demands too much

of a thinker" (p.49). According to the determinable-determinate relation, one must have some more determinate thoughts by virtue of which one is thinking the determinable thought. Woodling (2021) claims that even we can circumvent Grzankowski's criticism, Kriegel's approach will not save adverbialism because this approach takes for granted that "the terms fused together in the adverbial modifiers can be understood in isolation and not as the novel expressions they form" (Woodling 2021, p. 321). Another notable adverbialism is from Chirimuuta (2015). She proposes a theory called color adverbialism, claiming that Colors are "ways stimuli appear to certain kinds of individuals" and "ways that individuals perceive certain kinds of stimuli" (p.142). For a criticism of her view, see Gert 2017 pp.139-186.

5. The terminologies are form D'Ambrosio (D'Ambrosio 2019, pp. 418-419). The latter is also called the identity problem (Tye 1984a, p. 199) or the entailment problem.
6. The original version owes the credit to Jackson. For detailed demonstration, see Jackson 1975, p. 129-135.
7. I begin with hallucination, because adverbialism applies to all perceptual experiences and the form of adverbial analysis is the same no matter what kind of experience there is.
8. For details, see Sellars 1975, pp. 156-160; Tye 1975, pp. 138-143.
9. Another criticism of Tye can be found in Banick 2018, pp. 671-673.
10. "Coin" is an abbreviation of "coincidence". This part is a simplification of Tye's account. For details, see Tye 1984a, pp. 210-223.
11. Here "R" and "G" is the operator "redly" and "greenly" ("T" and "Sq" operate in the same way); "Hs" is the regimentation of "S hallucinates".
12. D'Ambrosio claimed that his approach can save adverbialism from the Many-Relations Problem (D'Ambrosio 2019, pp.434-436).
13. This part is a simplification of D'Ambrosio's account, for technical details, see D'Ambrosio 2019 pp. 427-438 and D'Ambrosio 2021, pp. 817-827.
14. An inference pattern holds within the complement of an intensional verb. For example, John seeks an F → John seeks a G, where F ⊆ G (D'Ambrosio 2019, p. 427).
15. Another criticism on D'Ambrosio can be seen in Gert 2021, pp. 702-708, who claimed that D'Ambrosio's view incurred invalid references, such as "if someone hallucinates a bottle of aspirin then they hallucinate a bottle of acetylsalicylic acid" (*ibid.*, 703).
16. A 'veridical hallucination' is a hallucination that happens to correspond with reality. A case of veridical hallucination can be found in Tye 1982, pp. 320-321.

17. [C] is akin to D'Ambrosio's characterization, which functions as the informative adjunction to capture the relation between certain event and its casual explanation. I use a single pair of square brackets instead of double to be distinguished from D'Ambrosio.
18. See previous discussion in 1.3.
19. I would like to thank the anonymous reviewer and Dr. David Pitt for their thoughtful comments which helped to significantly improve the paper.

BIBLIOGRAPHY

Banick, Kyle. (2018) "How to be an adverbialist about phenomenal intentionality," *Synthese* 1, pp. 661-686

Butchvarov, Panayot. (1980) "Adverbial theories of consciousness," *Midwest Studies in Philosophy* 1, pp. 261-80

Casullo, Albert. (1983) "Adverbial theories of sensing and the many-property problem," *Philosophical Studies* 44, pp. 143-160

Chirimuuta, Mazviita. (2015) *Outside Color: Perceptual Science and the Puzzle of Color in Philosophy* (Cambridge, Mass.: The MIT Press)

Chisholm, Roderick M. (1957) *Perceiving: A Philosophical Study* (Cornell University Press).

Corti, Luca. (2022) "Demystifying the myth of sensation: Wilfrid Sellars' adverbialism reconsidered," *Synthese* 2, pp. 1-21

Crane, Tim & French, Craig. (2021) "The Problem of Perception," *Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/perception-problem/>

Crane, Tim & Grzankowski, Alex. (2022) "The Significance of the Many Property Problem," *Phenomenology and Mind* 22, pp. 170-175

Dinges, Alexander. (2015) "The Many-relations Problem for Adverbialism," *Analysis* 2, pp. 231-237

Ducasse, Curt John. (1969) "Moore's refutation of idealism," in: *Truth, Knowledge and Causation*, pp. 192-217 (New York: Routledge)

D'Ambrosio, Justin. (2019) "A New Perceptual Adverbialism," *Journal of Philosophy* 8, pp.

413-446

D'Ambrosio, Justin. (2021) "The many-property problem is your problem, too," *Philosophical Studies* 3, pp. 811-832

Fish, William. (2021, 2nd) *Philosophy of Perception: A Contemporary Introduction* (Routledge)

Gert, Joshua. (2017) "Outside Color from Just Outside," *Philosophy and Phenomenological Research* 1, pp. 223-228

Gert, Joshua. (2021) "Adverbialism and objects," *Philosophical Studies* 2, pp. 699-710

Grice, H. P. & White, Alan R. (1961) "The Causal Theory of Perception," *Aristotelian Society Supplementary Volume* 1, pp. 121-168

Grzankowski, Alex. (2018) "The determinable–determinate relation cannot save adverbialism," *Analysis* 1, pp. 45-Jackson, Frank. (1975) "Symposium: The adverbial theory of perception. On the adverbial analysis of visual experience," *Metaphilosophy* 2, pp. 127–135

Kriegel, Uriah. (2008) "The dispensability of (merely) intentional objects," *Philosophical Studies* 1, pp. 79-95

Pautz, Adam. (2017) "Experiences are representations: An empirical argument," in: B. Nanay (Eds), *Current Controversies in philosophy of perception*, pp. 23-42 (New York and London: Routledge)

Sellars, Wilfrid. (1968) *Science and Metaphysics: Variations on Kantian Themes* (New York: Humanities P.)

Sellars, Wilfrid. (1975) "The adverbial theory of the objects of sensation," *Metaphilosophy* 6, pp. 144-160

Tye, Michael. (1975) "The adverbial theory: A defence of Sellars against Jackson," *Metaphilosophy* 6, pp. 136-143

Tye, Michael. (1982) "A causal analysis of seeing," *Philosophy and Phenomenological Research* 3, pp. 311-325

Tye, Michael. (1984a) "The adverbial approach to visual experience," *Philosophical Review* 93, pp. 195-226

Tye, Michael. (1984b) "Pain and the adverbial theory," *American Philosophical Quarterly* 4,

pp. 319-328

Van Steenburgh, E. W.. (1987) “Adverbial sensing,” *Mind* 96, pp. 376-380

Woodling, Casey. (2021) “Adverbialism, the many-property problem, and inference: reply to Grzankowski,” *Philosophical Explorations* 3, pp. 312-324

NO MORE (INTENTIONAL) BULLSHIT

Thomas Takasugi

SECTION 1: INTRODUCTION

Bullshit is everywhere and only seems to be getting more common, not only in everyday speech but in business, medicine, science, psychology, philosophy, and certainly politics. And because we frequently engage with the term, accusing and being accused of producing, spreading, or believing bullshit, a functional account of bullshit stands to both inform further academic investigation into the concept and help us identify, understand, and diagnose bullshit's broad presence in our own lives. Philosophical inquiry has so far developed two accounts of bullshit: intentional theories of bullshit that focus on the intentional act of bullshitting (Frankfurt 2005, Fallis & Stokke 2017, Kenyon & Saul 2022, Easwaran 2023), and non-intentional theories of bullshit that attempt to account for bullshit independently of its utterer (Cohen 2012, Wreen 2013). Although intentional accounts have so far dominated philosophical inquiry, I argue there are strong reasons to reject them as accounts of bullshit. Accordingly, the position I propose is a non-intentional account of bullshit that defines utterances that are both noxious and either irrelevant, fallacious, or conjecture as bullshit.

In Section 2, I discuss our colloquial, folk usage of bullshit and why having a concept matters. Then in Section 3, I introduce intentional accounts of bullshit before arguing why these accounts are fundamentally unapt to account for bullshit. In Section 4, I introduce and discuss two notable developments towards a non-intentional theory of bullshit by Cohen and Wreen. Finally, in Section 5, I present a new non-intentional account of bullshit that I argue is preferable to current alternatives.

SECTION 2: BULLSHIT

The subject of this paper is on what bullshit is, i.e., what we mean when we call something (e.g., an utterance) bullshit. I will be discussing only literal uses of 'bullshit' (i.e., uses of 'bullshit' that refer to bullshit); uses that do not necessarily purport to identify bullshit (e.g., emotive or hyperbolic uses[1]) are not relevant to the paper. Additionally, the methodology of this paper will appeal to intuitive use 'bullshit' as a folk term in order to derive a definition of bullshit. While

some may argue that bullshit is in the class of concepts including games for which conceptual analysis serves no practical purpose (I don't need to know what 'game' means in order to use it exactly as I should), I find these claims unconvincing because there *is* in fact at least one practical problem: People suck at identifying bullshit. Certainly, it's appealing to say we know bullshit when we see it, but actually seeing it—that is, correctly identifying it when it occurs—is not trivial and a failure to detect bullshit is often significantly harmful. Moreover, without a criterion to evaluate whether an utterance is bullshit, disagreement over alleged bullshit is liable to end in stalemate. A proper concept of bullshit, even if not all-encompassing of folk usage, doubles as a field guide in how to identify it by explicating what bullshit is. In Section 5, I present an account that I argue succeeds in this task.

Although competent folk usage does not entail understanding, when we call something bullshit our utterance is undoubtedly meaningful. It then seems appropriate to begin by discussing colloquial usage of 'bullshit'. Consider the following sentence:

(a) "Philosophy is bullshit."

I leave the construction of what sort of philosophy would prompt a response above to the imagination of the reader. As an initial try, one could give the meaning of 'bullshit' in (a) via synonym: Horseshit, hogwash, humbug, nonsense, bunk, baloney, drivel; the excrement of a bull. But this only shifts the challenge to defining another word that may or may not contain the concept of bullshit. Alternatively, one could look to a dictionary-like decomposition to help us articulate the concept. The OED defines 'bullshit' (noun) as "stupid or untrue talk or writing; nonsense," but this definition is too vague. If stupid utterances (whatever those are) suffice for bullshit, it would seem that we should *always* say that someone lacking in intelligence but putting forth a good faith attempt to contribute is bullshitting, but this is clearly false. If untrue utterances constitute bullshit, we are forced to call theories made in good faith but later deemed false (e.g., vitalism, luminiferous aether, Newtonian physics, etc.) bullshit.[2] Furthermore, although I think the OED definition under discussion is incorrect, I have no inclination to call it bullshit. 'Nonsense' seems the most plausible, but there are still good reasons to find it unsatisfactory. It may simply be that the pejorative element of bullshit should not be conflated with the more neutral sense of 'nonsense'. We must also be careful in how we define the slippery term 'nonsense' or risk unintentionally categorizing a great

deal of art as bullshit (and most art is not bullshit). But even if something's being nonsense could be called bullshit, it would be satisfactory only for a minority of uses. Nonsense, falsity, stupidity, or whatever other singular property cannot be a sufficient account of bullshit. Nonetheless, such terms are undoubtedly relevant in some capacity to the use of 'bullshit' in (a). On the other hand, consider an utterance of (b) where 'bullshit' is used as a verb:

(b) "Don't bullshit me."

The use of 'bullshit' here is referring to the act of bullshitting, not necessarily bullshit itself. The relationship between bullshit and bullshitting is disputed and two contrasting interpretations have emerged in the philosophical discussion of bullshit. Call these interpretations intentional and non-intentional accounts of bullshit. Intentional accounts of bullshit assert that an utterance is bullshit if and only if it is uttered by somebody who is bullshitting; non-intentional accounts of bullshit define bullshit via some set of criteria and maintain that someone who aims to utter bullshit is bullshitting. So, under an intentional account, the usage of 'bullshit' in (a) asserts that philosophy is bullshit *because* anybody who produces philosophy is bullshitting. In the next section I discuss intentional accounts in more detail and why I find them unconvincing as accounts of bullshit.

SECTION 3: INTENTIONAL BULLSHIT

In his 1986 essay *On Bullshit*[3], Henry Frankfurt offers an intentional account of bullshit. Although Frankfurt's discussion of bullshit resists a disciplined definition, let us say that a speaker is Frankfurt-bullshitting if they speak with indifference towards the truth value of their assertion.[4] It is an intentional account of bullshit because of its dependence on a speaker's particular state of mind satisfying the condition of indifference. Strictly, Frankfurt's theory is then an account of the act of *bullshitting*, however the theory still claims to account for bullshit: An utterance is bullshit if the *original* utterer is Frankfurt-bullshitting.[5] The added stipulation of original utterer (i.e., author) is crucial not only to Frankfurt's theory but to any intentional account as it allows intentional-bullshit to be spread among sincere individuals, provided the original author of the assertion was bullshitting. I will return to this point later in the section.

Frankfurt supplies a few examples to articulate his account, the most notable of which is the Fourth of July Orator.[6] Suppose an orator "goes on bombastically about 'our great and blessed

country, whose Founding Fathers under divine guidance created a new beginning for mankind'" (Frankfurt 2005, p. 16). Frankfurt's account maintains that since the orator doesn't actually care at all whether what she says is true so long as her assertions make her perceived by her audience in the way she wants (in this case, as a patriot), she is bullshitting. The conclusion that the orator is bullshitting certainly seems right, however Cohen (2012) and Kenyon & Saul (2022) argue that while the orator may not herself care whether what she asserts is true or false, she must care whether her audience believes that *she* believes what she says is true in order to achieve her intended enterprise of being perceived as a patriot. And since the orator must be indifferent about the truth value of her assertions in *any* capacity, Frankfurt's account cannot assert that the orator is bullshitting.

Although Frankfurt's account appears unsatisfactory as an account of bullshit, some have attempted to extend, clarify, or amend his account (Easwaran 2023, Kenyon & Saul 2022, Whitcomb 2023), while others have proposed alternative intentional accounts in its place (Fallis & Stokke 2017). Although these accounts resolve various objections to Frankfurt's account, they all share the common quality of, as intentional accounts, being based in a speaker's intention upon making an utterance. As I will argue for the remainder of this section, this foundation leaves us with good reason to reject intentional theories as accounts of bullshit.

Since intentional accounts disregard any context or content of the utterance, considering instead only the intention of the speaker, it is simple to construct a case where a speaker utters something true and relevant—neither of which relate to the speaker's intention—that nonetheless qualifies as bullshit:

> A high school biology teacher is required to teach a unit on evolution. They communicate all the relevant information about natural selection, phylogenetic records, etc., and tell their students statements like "humans and titi monkeys share a common ancestor." But the teacher is actually a hardline creationist and thinks everything they teach is false, worthless, and irrelevant to the real explanation of life.

By intentional accounts, the teacher is bullshitting their students. So, if the intentional account of bullshit is correct, then the teacher's true statements about evolution, uttered in a

relevant, scientific context, are bullshit. Moreover, this sort of argument may be run for *any* utterance (e.g., "2+2=4") in *any* context (e.g., a math classroom).

A dependence on intention also leads to an unintuitive epistemic limitation. Frankfurt claims that a bullshit utterance can be spread by sincere speakers if and only if the author was bullshitting, and I take this mechanism of transmission to be assumed in subsequent intentional accounts.[7] So, for an assertion to be bullshit while the utterer is sincere, there must be a causal chain of sincere usage that traces back to the original assertion made by a bullshitting author, if it originates from a single author at all. Even if we suppose some lone author exists, how are we to say that the utterer didn't truly believe in whatever we're now calling bullshit? Under intentional accounts, the strongest claim we can make about utterances made by causally distant authors is that they are *probably* bullshit on the basis that the original author was *probably* bullshitting. But this simply seems false. When we assert that something is bullshit, we really assert that it's bullshit. (And it doesn't seem like I need to know anything about the intentions of whoever authored theories of scientific racism to call it bullshit.) In any case, such a leap is not necessary: According to non-intentional accounts of bullshit (including the one I propose in Section 5), while the original author(s) of such utterances may indeed *probably* be bullshitting, the utterances under discussion are nonetheless plain-and-simple bullshit.

Intentional accounts also face the challenge of accounting for situations where it seems no intention is present, yet the utterance is nonetheless bullshit. Consider the following case where no bullshitting entity exists:

> A student with strep throat attends class anyways. Her classmates voice their concerns that she is contagious and should not be present, or at a minimum should be wearing a mask, however she assures her peers that she looked into it online and saw that strep throat is no longer contagious after the first day of infection. (In reality, this is false and the student had misread what was really the contagious period *after* taking antibiotics (which she hasn't taken).) She is sincere, so the other students trust her and continue on with class.

Medical misinformation, especially of the directly harmful kind above, is surely bullshit. Yet under an intentional account, we should not say that the utterance is bullshit because the student clearly is not bullshitting and there is no original author that we might attribute the

utterance to—the website she read provided correct information. At worst, she is acting irresponsibly and without due diligence, but she isn't bullshitting. Further, the prevalence of utterances without intention only stands to grow as generative artificial intelligence sees wider and more general use, yet it is unclear how we can generalize the creators' intentions to such broadly applicable tools.

If intentional accounts of bullshit are to be taken seriously as accounts of bullshit (and not merely accounts of the act of bullshitting wholly disconnected from bullshit itself), they must address these challenges. I do not think they can succeed. In the next section, I consider accounts of bullshit that avoid these problems by making no appeal to intention.

SECTION 4: NON-INTENTIONAL BULLSHIT

In response to Frankfurt's theory of bullshit, G. A. Cohen (2012) [8] offers an alternative approach: investigate bullshit itself, regardless of its utterer's mental state. I call approaches of this kind non-intentional accounts of bullshit because an utterer's intentional state bears no relation to whether or not an utterance is bullshit. Non-intentional accounts also stipulate a different relationship between bullshit and bullshitting: A speaker is bullshitting if and only if they *aim* to produce bullshit, and a speaker is a bullshitter if and only if they are *disposed* to bullshitting. That non-intentional accounts stipulate no necessary connection between the intentional act of bullshitting and utterances constituting bullshit immediately yields intuitive results that align with our folk conception of bullshit: Someone who is bullshitting might inadvertently say something true and relevant, while someone who is speaking in good faith may inadvertently utter bullshit (these points were highlighted in the previous section). Furthermore, without appeal to intention, problems stemming from our access to (or the presence of) an author's intention do not arise. In short, the arguments I gave against intentional accounts do not apply to non-intentional ones, and this gives us good reason to pursue a non-intentional account.

Cohen limits his discussion to just one kind of bullshit that he calls nonsense, defined as "unclarifiable unclarity" (p. 333). He asserts that an utterance is unclarifiable unclarity if it cannot be made clear, and while he does not articulate what "clear" means, he offers that a statement is sufficiently unclear if its plausibility is unchanged with the addition of a negation sign. Cohen is also wary of categorizing poetry or other valued kinds of text as bullshit and blocks their inclusion

with the added condition that bullshit must be unclarifiable *and* lack a "virtue of suggestiveness" that makes interacting with a particular thing worthwhile (p. 334)

In accounting only for a small subset of a subset of bullshit (i.e., unclarifiable unclarity), Cohen's timid theory highlights a significant challenge facing non-intentional accounts: If not careful while introducing contextual and content-based criteria, one risks categorizing too many things as bullshit. Michael Wreen (2013) proposes a non-intentional account of bullshit that is much stronger than Cohen's and also appears to address this challenge. He claims that *all* bullshit is noxious nonsense, where, relative to a discourse, 'noxious' utterances are worthless, harmful, or of negative value, and a nonsense utterance is one that

(1) doesn't fit in with or contribute to an understanding or resolution of the issue at hand,
(2) doesn't advance knowledge of the phenomenon under consideration, or
(3) doesn't help to solve the problem that's being addressed (Wreen 2013, pp. 111-112).

Wreen's theory results in a much more robust non-intentional account that is also desirably restrained due to the additional, widely applicable condition of noxiousness. A poem, even *if* it is truly undecipherable nonsense, is not bullshit under Wreen's account if we take its contents to be noxious. Colloquially, we might liken noxiousness to toxicity as it occurs in phrases like "toxic masculinity." An engineer uttering technobabble in an important technical meeting is uttering bullshit, while an actor portraying a Star Trek character who utters equally nonsensical technobabble is not uttering bullshit; the difference is wholly in the harm or lack of value such utterances contribute to the relevant discourse. (One might think that the difference lies in Star Trek's status as fiction; I would conjecture that there is no context safe from bullshit's reach, but this is not a consequential point.) The condition of noxiousness is surely vague, but vagueness here should be welcome: Noxiousness may instantiate in ways that no theory is likely to satisfactorily capture, yet where it does occur it is apparent and demonstrable. Keeping noxiousness vague accommodates anything from acute noxiousness, as an immediate effect, to indirect noxiousness that is harmful in principle but does not immediately affect anyone or causes stochastic harm.

Wreen's view brings us much closer to a satisfactory non-intentional characterization of bullshit, but there is still a good deal of room for improvement. Particularly, the criteria (1)-(3) are too vague. And unlike the vagueness in noxiousness, the operative terms in (1)-(3)—

understanding, resolution, knowledge, and solution—are beholden to a wide variety of meanings that may be demonstrable only near the ultimate level of a discourse. For example, it is unclear whether Wreen's 'nonsense' captures Cohen's unclarifiable unclarity. Suppose a charlatan sells somebody some life advice that's total unclarifiable unclarity by Cohen's standard (perhaps the utterance isn't even syntactically coherent), yet the recipient, for whatever reason, finds personal meaning from the nonsense. Both Cohen's account *and* intentional accounts would call the charlatan's utterance bullshit, and this seems like the intuitive answer as well. Yet whether Wreen's account concurs depends on how the criteria (1)-(3) are interpreted, and it is easy to draft plausible interpretations that conflict. In practice, disagreement over whether an utterance is bullshit is to be expected and providing a framework for this disagreement is a primary goal of the paper. However, in disputes over whether a claim is bullshit, the definition of bullshit itself should be kept fixed: Opportunity for equivocation should be avoided, if possible.

The purpose of this analysis is to give insight into what our concept of bullshit really is and I am unconvinced that Wreen's account satisfies this pragmatic criterion, especially in discursive climates where felicitous conversation can rarely be presupposed. Yet despite the apparent flaws in Wreen's account, his is the most developed non-intentional account to date and I think his assertion that bullshit is noxious is correct. In the next section I propose a non-intentional account that develops this insight of Wreen's with a different set of accompanying criteria. This new account, I argue, is preferable.

SECTION 5: IRRELEVANT, FALLACIOUS, OR CONJECTURAL BULLSHIT

I propose the following definition of bullshit:

> (Bullshit) An utterance is bullshit if and only if it is both noxious *and* either irrelevant, fallacious, or conjecture.

This account can be seen as influenced by Cohen and a development of Wreen's. I take noxiousness to be a necessary component of bullshit discourse and expand Cohen's briefly suggested (but never meaningfully explored) categories of rubbish and irretrievable speculation that approximately correspond with fallaciousness and conjecture. As noxiousness was discussed

in the previous section, my plan in this section is to present the three terms of the disjunct, examine some applied instances of this account, then discuss why this account is preferable.

I will begin by discussing the conditions upon which an utterance is irrelevant. Wreen's (2013) 'nonsense' vaguely identifies the relation between bullshit and irrelevance, however Fallis & Stokke (2017) identify this relation much more directly and employ the question-under-discussion (QUD) framework introduced by Stalnaker to identify when an utterance is (intended to be) irrelevant. I think the QUD framework would be proper to apply here, however, a proper introduction of the QUD framework is beyond the scope of this paper. Moreover, I don't think it's necessary.

> (Irrelevance) An utterance *u* is irrelevant relevant to a topic of conversation *c* if and only if *u* is not a relevant response to *c*.

This may appear thin but I take it to be as substantial as need be. Like noxiousness, I take what is relevant in a discourse to be salient and demonstrable; irrelevant utterances are just those utterances that violate Grice's maxim of relation: Be relevant. Given a topic of conversation like who murdered Smith, reasonably relevant responses might include "The murderer is insane!" or "Are we sure there weren't two murders?" or "I don't see Jones anywhere;" an irrelevant response may resemble "Oops, I forgot to pay may taxes." One might object that this definition remains unclear because adversarial conversants may disagree on what the topic of conversation is or what constitutes a relevant response. These disagreements undoubtedly occur, but they are disagreements over what the topic is or what is relevant, not what a topic is or what relevance means; once again, the concepts are not under dispute and so these disagreements are acceptable.

Before I stipulate what I mean by "fallacious," I want to briefly address how fallacy is often misunderstood. Fallacies, especially informal fallacies like ad hominem, slippery slope, or strawman, are often notorious not for their rational force in discussion, but for their hasty, liberal, and often unwarranted use in what might aptly be called pseudo-rational contexts. The mere observation of reasoning that resembles the form of a fallacy is not sufficient grounds to then conclude that the reasoning is fallacious. Observation of fallacy is only a heuristic indicator that a failure of reasoning may be occurring; due diligence requires us to inquire further to find out *where* the failure has occurred. With this in mind, I propose the following:

(Fallacious) An utterance *u* is fallacious if it is false or based on a bad argument.

An argument is a set of premises and a conclusion, and arguments may be formal (e.g. *modus ponens*: $P \supset Q, P \vDash Q$) or informal (e.g. inference to the best explanation). If it is possible for the argument to have true premises and a false conclusion, the argument is invalid; if a valid argument has a false premise, it is unsound. Formal arguments are bad if they are either invalid or unsound. Any argument (both formal and informal) is bad if it does not contextually substantiate an utterance *u*. A sound formal argument that's misapplied within a context is a bad argument, even if the utterer believes it to be relevant and informative. An informal argument might be bad if it appeals to insufficient empirical evidence. The markers of a bad argument in practice may be difficult to pinpoint in practice, but this is why critical reasoning skills are indispensable.

What I call fallacious Sagan (1995) might call baloney. Conveniently, then, Sagan's "baloney detection kit"—a collection of valuable heuristic methods for practicing inquiry and distinguishing good arguments from the bad—then doubles as a bullshit detection kit, provided the condition of noxiousness is also met. Knowledge of these sorts of tools, insofar as that knowledge serves as a heuristic indicator of deficient reasoning or unacceptable connection to evidence, is directly related to one's ability to identify bullshit. Education in critical reasoning is a natural step towards understanding and articulating *why* an utterance *is* bullshit. If this much is true, even if philosophy were a culprit source of bullshit, it may yet be simultaneously true that philosophy is among the best equipped disciplines to cultivate the skills necessary to combat it.

The claim that fallaciousness (as I have stipulated it) encompasses all false utterances may raise concerns about how my account differentiates bullshit from lies. The relationship between lies and bullshit has been a significant topic of concern for intentional accounts. Recent works have accepted that many lies are also bullshit but deny that the concepts are coextensive (i.e., not all lies are bullshit and vice-versa). However, while lies are indeed necessarily intentional, bullshit (under non-intentional accounts) is not. Therefore, we may simply characterize the conceptual and psychological significance of a lie by its intention—namely, the intention to cause an audience to uptake a belief that the utterer thinks is false. Bullshit and lies are clearly not coextensive because many lies do not satisfy the condition of noxiousness, and one may lie by uttering a true statement one believes to be false. Where they do coinstantiate, the term we use indicates the utterance's

more significant feature. Say we have some utterance *u* that is both bullshit and a lie. If the egregious falsity of the claim is what we wish to emphasize, we call *u* bullshit, and if the duplicitous intention is our target, we call *u* a lie. And if we can't decide, we may simply call *u* a bullshit lie. *Even if* bullshit and lies were coextensive, it seems we would still be fully capable of disambiguating and studying the concepts. Perhaps it is merely the advantage of a non-intentional framework, but I am unconvinced the alleged problem of bullshit-lie overlap is a problem at all.

The final category I propose is conjecture. What I am calling conjecture differs from fallacy in that while an utterance may be fallacious for its basis on an argument made bad by lack of acceptable evidence, conjectural utterances lack any argument at all. Additionally, I offer that if an utterance cannot be known true or false when it is made, it is conjecture. This resembles Cohen's negation test for unclarity. For example, Bertrand Russell's claim that a teapot orbits the sun is conjecture because although it is physically conceivable that we could detect such an object in such a position, it is certainly unknowably true or false by any means available to Russell in his lifetime.[9] The same holds if we claim that there is not a teapot orbiting the sun. Let conjecture be defined as the following:

> (Conjecture) An utterance *u* is conjecture if and only if *u* can't be known true or false or contains an empirical claim that can't be substantiated with sufficient evidence.

The standard of sufficient evidence may vary according to the strength of the corresponding utterance, and whether an empirical claim is substantiated by sufficient evidence is not always immediately knowable. Further inquiry may be required in order to determine whether an utterance is conjecture, but this is hardly a detriment—it is already the case that we should strive to ask follow-up questions during any inquiry into bullshit. In cases where no follow-up questions may be asked (e.g., the author is dead), we may be compelled to work with the information available to us. While this is an unfortunate epistemic limitation, I suspect it will remain until we learn to speak with the dead.

Now that the components of my account have been defined, I will consider how this account responds to some cases of bullshit utterances. Fallis & Stokke consider the following story of a lying umbrella salesperson as an example of bullshit:

> Louise wants to sell Tom an umbrella. She knows that Tom is going to Chicago. Even though Louise believes the opposite, she invents the story that it is always raining in Chicago at that time of year, and she tells Tom that in order to make him buy the umbrella. (p. 304)

Conjuring false information to get one's way and make someone else waste their money seems to readily satisfy the condition of noxiousness, and because Louise's story is false, it is fallacious (as I have defined the term). Because Louise is lying, it is also a lie. Whether we call her story bullshit or a lie depends on what quality we wish to emphasize, but its status as bullshit and a lie is clear. Alternatively, suppose Louise hasn't the faintest idea what the weather in Chicago is like this time of year. Her utterance would then be both noxious fallacy *and* noxious conjecture, since not only is her claim false, but she doesn't have any basis to know that fact, even if it were true. That Louise has uttered bullshit may only be known to her in the moment, but Tom may learn that she is saying bullshit with a simple internet search. He needn't obtain a true report of her internal state to make that conclusion.

Recall the case of the student with strep throat that I discussed in Section 3: A student catches strep throat and tells her classmates that she isn't contagious. She believes that she is telling the truth but is actually spreading medical misinformation. As I mentioned in Section 4, an advantage of my proposal as a non-intentional theory is that bullshit may be properly labeled even when no intentionally bullshitting entity exists. The student's misinformation is obviously harmful to the health of her peers. It's also fallacious due to her appeal to false evidence that came not from the website that gave sound medical advice, but her own carelessness. Her claim is bullshit, but again, we should not say that the student is bullshitting because she did not aim to utter bullshit.

As a final example, I want to discuss the well-known peddler of pseudoscience, Deepak Chopra. Consider the following social media post by Chopra:

(c) "The present moment is a conspiracy of the total universe."[10]

The harmfulness of (c) may be debated, however I appeal to the strong intuition that utter nonsense asserted as serious insight is always noxious to fruitful discourse. It is also immediately evident that (c) *at least* satisfies the condition of conjecture—it is not at all clear on what grounds

we could call (c) true or false, or if doing so is even possible at all. In a scientific context, (c) would also constitute irrelevance. Without taking any stance on whether Chopra holds sincere intentions, claims such as (c) are bullshit.

Following recent psychological research there is incentive to develop a better definition of bullshit. Pennycook et al. (2015) are interested in "the factors that predispose one to become or to resist becoming a *bullshittee*" (p. 550) and propose the Bullshit Receptivity (BSR) scale as an empirical metric of an individual's sensitivity to bullshit utterances. In that paper, Pennycook et al. investigate what they call "pseudo-profound bullshit" and utilize Frankfurt's definition of bullshit that defines bullshit utterances as utterances made with indifference to the truth. This operational definition is carried into subsequent studies of bullshit receptivity (Sterling et al. 2016, Pennycook & Rand 2020, Evans et al. 2020; but see Pfattheicher & Schindler 2016 who operationalize "pseudo-profound bullshit" without intentional language). Wilson (2023) observes a significant problem in the widespread use of Frankfurt's account of bullshit: It is "explanatorily inert" (p. 5). The criterion of Frankfurtian bullshit is trivially satisfied by a study's utterance materials being either randomly generated or selected for their vague language, and in neither case for their truth or falsity. More critically, the BSR scale is not designed to measure a subject's receptivity to the speaker's intentions,[11] and so it is unclear how Frankfurtian bullshit (or any intentional account, for that matter) is relevant. Instead, Wilson argues that the BSR scale should be framed in terms of Cohen's definition of nonsense as unclarifiable unclarity.

Pennycook et al. are aware that pseudo-profound bullshit is only one of many types of bullshit and call for further research on bullshit receptivity to "be clear about the type of bullshit they are investigating" (Pennycook et al. 2015, p. 561). This call is reiterated in Sterling et al. (2016). The non-intentional account of bullshit I have given addresses these concerns: My account maintains the utility Wilson assigns to Cohen's account via conjecture and also suggests two additional types of bullshit. However, unless pseudo-profound statements are taken to be noxious as a rule, the BSR scale, because it does not consider research participant opinions on the noxiousness of given utterances, tracks pseudo-profound nonsense, not bullshit. This is largely a pedantic claim that has no bearing on the validity or significance of the empirical results of the studies discussed, however it may be valuable in future BSR studies to control for the perceived value of pseudo-profound statements, especially when investigating their correlation with political support.

Finally, I have so far discussed bullshit only as it pertains to utterances, however I imagine the account I have given can easily be applied, *mutatis mutandis*, to non-utterance bullshit (e.g., representations, actions, institutions, laws, etc.) as well. David Graeber's (2013) famous "bullshit jobs" seem to exemplify noxious irrelevance. Similarly, we might call a law bullshit if it obstructs justice and fails to address what it was designed to legislate or is predicated on a bad argument. This discussion, while promising, unfortunately lies outside the scope of this paper.

SECTION 6: CONCLUSION

Intentional accounts of bullshit, which still currently dominate academic discourse on the subject, are deeply flawed as accounts of bullshit and, at best, can be taken to address the intentional act of bullshitting and nothing more. The account I have offered is a non-intentional account that articulates three types of bullshit—irrelevant, fallacious, and conjectural—unified under the necessary condition of noxiousness. Simply put: an utterance is noxious if it is worthless, harmful, or of negative value; irrelevant if it lacks sufficient relation to a discourse; fallacious if it is based on unsound reasoning or faulty evidence; and conjecture if it has no basis or cannot be known to be true or false. I argue that this account is preferable to alternative non-intentional accounts both for its improved clarity and explanatory value.

In discussion of a subject such as bullshit, it is crucial to remember that our common enemy is bullshit itself—the noxious content that reasonable minds find repugnant—not those who unintentionally produce and spread it. So, when we encounter somebody stuck in the mud, our first response shouldn't just be to put a sign next to them reading "bullshitter" and move on; help them and share the skills that will prevent them from getting stuck again. This is much easier said than done, but it's a step we must all acknowledge as necessary if we ever hope for greener, bullshit-free pastures.

Notes

[1] Emotive uses are typically expressions of dissatisfaction, disapproval, or outrage, and are functionally interchangeable with any other word, or even a strong grunt; hyperbolic uses, insofar as they are not meant literally, need not literally refer to bullshit.

[2] On the other hand, I think a proper account of bullshit should be able to declare some (if not all) pseudoscience as bullshit, even if the advocates of such pseudoscience are sincere.

[3] Harry Frankfurt originally published *On Bullshit* as an essay in 1986, however citations in this paper will refer to the book of the same name published in 2005. The content is the same.

[4] Frankfurt asserts an additional condition in *On Bullshit* of intentional misrepresentation about the fact that one is indifferent to an utterance's truth, however its inclusion is not necessary to demonstrate the problem of Frankfurt's account. Moreover, this condition is not reasserted in Frankfurt's reply to Cohen (2002).

[5] This is asserted by Frankfurt in his reply to Cohen (2002).

[6] In *On Bullshit*, Frankfurt notes that the 4th of July Orator example originates from Max Black's 1985 discussion of humbug and Frankfurt uses the orator to demonstrate Black's account of humbug, not bullshit. Despite this, the Orator example has been featured prominently in the literature as an example of bullshit (Kenyon & Saul 2022, Easwaran (2023) Fallis & Stokke (2017) Cohen (2012), Whitcomb (forthcoming)), most likely because it is nonetheless the most straightforward example of Frankfurt's bullshit in the text. With this justification in mind, I will also discuss it as an example of Frankfurt's concept of bullshit.

[7] Any serious intentional account should contain a mechanism of bullshit transmission. Charitably, if no mechanism of transmission is explicitly given by an intentional account then one should be assigned to it. Since I am aware of no other such mechanism given by an intentional account and I cannot think of one that would be preferable to Frankfurt's, I assume Frankfurt's mechanism applies to all intentional accounts.

[8] Originally published in 2002.

[9] Russell's conjecture of such a teapot is obviously an argument by *reductio*, and I'm sure he would call any sincere claim of an entity such as the orbiting unobservable teapot bullshit. Insofar as that sort of conjecture is also harmful to a felicitous discourse, my account would agree.

[10] Found via wisdomofchopra.com. Original source: https://twitter.com/DeepakChopra/status/210881940570636288

[11] In the four studies done in Pennycook et al. 2015, research participants were given no explicit context that might provide information about the author of a statement's intention.

Bibliography

Easwaran, Kenny. (2023) “Bullshit activities,” Analytic Philosophy, 00, pp. 1-23

Kenyon, Tim & Saul, Jennifer. (2022) “Bald-faced bullshit and authoritarian political speech: Making sense of Johnson and Trump,” in: Laurence Horn (ed.), From Lying to Perjury: Linguistic and Legal Perspectives on Lies and Other Falsehoods. De Gruyter Mouton. pp. 165-194

Fallis, Don & Stokke, Andreas. (2017) “Bullshitting, Lying, and Indifference toward Truth,” Ergo: An Open Access Journal of Philosophy 4, pp. 277-309

Whitcomb, Dennis. (2023) “Bullshit Questions,” Analysis, 83, pp. 299-304

Wilson, Jonathan. (2023) “Rethinking Bullshit Receptivity,” Review of Philosophy and Psychology

Cohen, G. A. (2012) “Chapter 5. Complete Bullshit,” in: *Finding Oneself in the Other*. Princeton University Press. pp. 94-114.

Frankfurt, Harry G. (2002) "Reply to G. A. Cohen", In: S. Buss & L. Overton (Eds), *The Contours of Agency: Essays on Themes from Harry Frankfurt*, pp. 340-344 (MIT Press)

Wreen, Michael. (2013) “A P.S. on B.S.: Some Remarks on Humbug and Bullshit,” Metaphilosophy 44 (1-2), pp. 105-115

Pennycook, Gordon et al. (2015) “On the Reception and Detection of Pseudo-Profound Bullshit.” *Judgment and Decision Making* 10.6, pp. 549–563

Pennycook, G., & Rand, D. G. (2020) “Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking,” *Journal of Personality*, *88*(2), pp. 185–200

Evans, A., Sleegers, W., & Mlakar, Z. (2020) “Individual differences in receptivity to scientific bullshit,” *Judgment and Decision Making*, *15*(3), pp. 401–412

Sterling, Joanna, John T. Jost, & Gordon Pennycook. (2016) “Are Neoliberals More Susceptible to Bullshit?” *Judgment and Decision Making* 11.4, pp. 352–360

Pfattheicher, Stefan, & Schindler, Simon. (2016) “Misperceiving Bullshit as Profound Is Associated with Favorable Views of Cruz, Rubio, Trump and Conservatism.” *PloS ONE* 11.4, e0153419

Frankfurt, Harry G. (2005) *On Bullshit*. Princeton, NJ: Princeton University Press

Sagan, Carl. (1995) *The Demon-Haunted World: Science as a Candle in the Dark*. 1st ed. New York: Random House

Black, Max. (1983) “The Prevalence of Humbug,” in: The Prevalence of Humbug and Other Essays, pp. 115–143 (Ithaca, N.Y.: Cornell University Press)

Graeber, David. (2013) "On the Phenomenon of Bullshit Jobs: A Work Rant". *Strike Magazine*. Archived from the original on August 7, 2018

AFFECTIVE EXHAUSTION IN THE AGE OF SEMIOCAPITALISM

Albert Cardenas

INTRODUCTION

Postmodernity is marked by the intensification of free-market capitalism. Production of capital is no longer dictated by materiality, but the immaterial – i.e., information, communication and discourse, affect, atmosphere, desire, signs, libidinal investments, etc. There, where the unrestrained expansion of the immaterial plane clashes with the limits of human cognitive abilities, is where the next frontier of production emerges - *semiocapitalism*. That is, the mutation of cognitive capacities into productive vessels of economic activity without directly materializing them, mediated by semiotic operators. Gone are the days of mass factory labor, and as such, labor being distinctively physical. It is not that exploitation of physical labor no longer occurs, nor that material production has ceased to occur. Rather, in the 21st century (and particularly in the Global North), we are witnessing a sharp increase in cognitive exploitation—the economic operationalization of human cognition, oriented towards the proliferation of sign-value that guides internalized drives for consumption and production of sign commodities. Marxist grammars of value can no longer adequately capture the power of the commodity as the hegemony of the sign reigns supreme over its function (use) and measurable labor required for its production (exchange). Thus, the Marxist grammars of alienation take on a different character—virtualized hell that degrades the psyche of cognitive workers. In a word: depression. The disappearance of sensuousness, imagination, sensibility, and the future itself—that which is imperative to enjoying life beyond the everyday death march towards continuous self-optimization. Reality too is disappearing—for it was Jean Baudrillard who predicted the impending arrival of a sign economy legitimized by models, or technologies of simulation, that stand in for the real. Today, Franco Berardi warns us of "neuro-totalitarianism," but it was Baudrillard who had long predicted that we would one day become our screens—the inescapability of "hyperreality" made possible by incessant flows of semiotic capital. Depression, anxiety, and docility become the primary modes of violence that postmodern subjects experience (if I can refer to one as such) – a relentless assault on the shared nervous system, or "psychosphere" that results in the fragmentation of human subjectivity. The rhythm of labor is now attuned to the domain of the infosphere, cyberspace, and mass media. Adapting one's mind to accelerating rhythms of semiotic production presents

limitations, in that, the mind can only consume, store, and produce enough semiotic content before it reaches critical overload, the point of excess that makes psychosomatic life ever prone to affective exhaustion—or worse, suicide. It is not solely an excess of semiotic stimuli that the mind must bear but an excess of that content such that its meaning implodes under its weight—the loss of meaning itself. Virtualization of life demands that we 'play the game'. It is play or lose, and to play the game is to simultaneously bear and evade the nearing certainty of psychosomatic violence.

Where state politics and radical fantasies either reproduce the status quo, or miss the revolutionary boat, I am interested in combating semiocapitalism at a symbolic level. This primarily involves recuperating waning psychic energy that thrusts subjects into debilitating modes of depression. I say debilitating because it is debilitating, first and foremost, to the subject's overall welfare and to hopes of forging social solidarity that is necessary for creating new modes of resistance against the flows of capital. Throughout this article, I will further elaborate on Berardi's theory of semiocapitalism, Jean Baudrillard's complementing of Berardi's formalization of semiocapitalism, and lastly, offer a few lines of flight away from a life of constant productivity.

SEMIOCAPITALISM PROPER

Semiocapitalism adjusts the cognitive capacities of the human to quick and intense rhythms of capital production, creating psychosomatic breakdowns as cognitive laborers are subjected to affective burnout. Accelerated production and consumption of semiotic content open one up to the weight of incessant information, vertigo from feeling overwhelmed, and barren signification (or loss of meaning). For example, a music streaming app that has algorithms with the ability to categorize music according to your 'mood' or 'vibe.' The idea that one's sensibilities can be manipulated towards economically productive ends means that one is never truly free from exploitative labor. One is always working, unbeknownst to them, submitting to a process of datafication that impersonally reduces flesh to code. In *the Uprising: on Finance and Poetry*, Berardi explains that because language becomes infected by financialization (or, profit motive), expanding digital networks upon which financial exchanges occur implies that direct profit is being acquired without the material production of commodities. Since sign-value is ultimately what is exchanged in this phase of capitalism, and simulation divorces signs from their base referents, "Financial signs have led to a parthenogenesis of value, creating money through money without the generative intervention of physical matter and muscular work" (Berardi 2012, p. 19).

Baudrillard understood this decades prior in *The System of Objects*, where material goods participate in a "hierarchy of signs" that inform mass consumerism. In other words, people don't buy an object per se, but rather the social relation the object carries by way of what it signifies. The right brand (heavily mediated by advertising imagery) allows consumers to purchase social status, or a particular perception like affluence. We have gone beyond the marketplace of material goods and have entered the marketplace of signs. In the following passage, I will identify a few examples to concretize the effects of semiocapitalism.

BARREN COMMUNICATION

Consider the expansive nature of digital networks that can facilitate information maximization and the overproduction of meaning. This is an excess of information that is available; infinite growth is achievable on the web. As the social body becomes heavily mediatized, communication becomes more digital and fractured and undergoes a process of a-signification - meaning the internal content of communication becomes meaningless. Baudrillard observes this in *Simulacra and Simulation*, with a greater focus on *The Ecstasy of Communication*. In particular, he posits that information becomes dissuasive in the new media age. As new modes of media acquire the capacity to communicate information in greater breadth, the meaning of that information significantly loses depth in meaning. As the human subject plunges deeper into the infosphere, the rate at which information is uploaded is too intense such that the mind cannot process or make sense of the information being consumed. The consumption of information then becomes more equivalent to being situated in the midst of a chaotic blur of signs that lack semblance; it all reflects a sameness, and no greater insight is extracted. The vastness of cyberspace is excitable in that it is ceaselessly neuro-stimulating—which is to say, being digitally plugged in requires the extraction of affective, libidinal, and intellectual energy. Berardi proposes that under the paradigm of semiocapitalism, one's labor becomes "cellularized"—or, bought as units of time. Essentially, one is always 'on call' as semiotic production nearly occurs throughout the day. The important distinction to draw between cellularized labor and industrial labor is that industrial labor occupies a space from which the body can remove itself. For example, one may clock into a workspace for a standard period of time and return home, or wherever, but maintaining a distance from the act of labor. One supposes that outside of the workspace is where time is freed up for the worker to engage in other activities that are fulfilling, pleasurable, etc. Cellularized labor, though, suggests that one is no longer limited

to a single workspace as the work environment is potentially everywhere where the worker is plugged in. As such, does one have any free time to decompress from the stress of labor? The working conditions of semiocapitalism imply that one is always working. In fact, when one believes they are engaging in leisurely activities, they are only simulating leisure under expanded definitions of labor. For example, my generation and those after (Internet natives) who have become accustomed to traversing social media as a casual form of leisure are still on the clock. Whether participating in discursive exchanges on message boards or consuming visual entertainment, it still falls under the domain of immaterial labor. Clicking a 'like' here and sharing content there contributes to the schema of semiotic production that encapsulates the infosphere. Nebulous, panoptic algorithms that track and collect digital activity tailor your user experience such that it is easier for you to consume enjoyable content. This is only manipulating desire, effect, etc., for further sign production and consumption. Accumulation of value within symbolic economies that directly create profit strip language and symbols of their meaning, and instead, are repurposed towards economic ends. This is most commonly observable in the deployment of hashtags, emojis, memes, and digital discourse, among others, that orient limited economies of attention toward those signs that propagate spectacles. Much like branding, if audio-visual content is promoted within a trending discursive framing in digital circulation, it can compel consumers to consume that particular object among and alongside other users. Human sensibility is at the mercy of semiotic production as the consumption of sign commodities is never-ending, entrenched in a perpetual merry-go-round of commodification. Communicative exchange becomes impossible as media technologies advance, precisely because language is already at risk of commodification. Consuming excess information and making sense of it is an arduous task that produces sensations of vertigo and disorientation from the stream of information that does not pause or decelerate.

It is worth noting the rise in particular trends and changes in behavioral patterns of consumption at the intersection of leisure and cellularized labor. Since time is always fleeting, one feels compelled to accelerate their modes of audio-visual consumption. Consider the language of multi-tasking - it is the active testing of one's neuroplasticity to maximize productive labor. There is not enough time to work, enjoy leisure, and vice versa. So, one performs both at once, which may look like performing several tasks at once. It is now possible to file taxes, submit assignments, send correspondence, read, and consume audio-visual entertainment from one screen—what a time to be alive! One must also multitask, however, if only to cope with the flows and rhythms of

semiotic production—like a survival strategy. But what does survival mean when it is disguised as psychosomatic adjustments—or, more accurately, discipline—in service of maximizing the accumulation of value? Furthermore, because there is a lack of time, more people have begun consuming audio-visual content at accelerated speeds because their attention economy is attuned to the rhythm of semiocapitalism. If it can be consumed, it must be instantaneous so as to keep up with demands of productivity. There is no time to truly enjoy experiences if they do not conform to the tempo of cognitive labor. Consequently, the subject is estranged from slow time and now desires a particular mode of consumption to be delivered at fiber optic speed. This affects not only consumers of audio-visual content but creators as well. Tristam Vivian Adams reminds us, "The accelerating demands of semiocapitalism that exploit our communicative and cognitive capabilities are reflected in syllabic speed. We are in the business of talking faster" (Adams 2018, p. 33). As such, the need for speed is commonly reflected in the delivery of speech on platforms such as YouTube or TikTok. This is less so an aesthetic pertaining to the delivery of speech than it really is about the tempos of semiocapitalism infecting speech patterns. Adams continues, "Hyper-spliced, energetically cut and chopped YouTube videos that omit any hesitance, pause or delay from the speaker are a growing online video aesthetic. Today our exposure is as much to technologically accelerated syllabic intensities, as it is to those within the human remit of syllabic delivery frequencies" (Adams 2018, p. 38). Because so much visual entertainment is mediated by technologies of audio-visual perfection, speech can be edited to mimic the speed at which semiocapital flows. It is the glorification of productive speech and the erasure of the organic; this becomes a dominant mode of expression projected onto the visual-sonic world, announcing itself as a desirable way of operationalizing the voice. I will elaborate on this notion of the voice in the following passage.

SEMIOTIC INFLATION OF EDUCATION

Elsewhere, university students, who are intellectual workers, never cease to work—partly due to neoliberal prescriptions of educational programming whereby the pursuit of academic excellence at all costs creates a semiotic inflation of values pertaining to GPAs, test scores, grades, etc. This begins in the primary phases of early education. Contemporary education practices demonstrate an urgency to strive toward academic excellence, chasing the image of the ideal student. Contemporary liberal narratives of economic progress are closely connected to securing a 'good

education'. That is, the value of education is reflected by its economic ends - namely, financial security. Universities and college readiness programs, among others, are guilty of creating a toxic will to productivity that generates competition among students. Technologies of consumerism are clever in gaslighting consumers because they teach them to repress their capacity to self-create. In other words, consumers are taught to repress themselves and map the fantasy of their ideal self onto objects of consumption and advertising, encouraging them to use their purchasing power to fulfill a lack. One can fulfill that lack through an overidentification with an object of consumption, or its sign-value. Regretfully, it is no different in the context of contemporary education. Why is one student sold on attending a university such as Harvard, Yale, or Berkeley, as opposed to their local state college? Universities, like any other object of consumption, operate within a hierarchy of signs that compete against each other for greater visibility, enabling them to remain within circulation on the semiotic market. Students also undergo a massive process of sign commodification, which occurs at a very basic level in all levels of instruction. For example, the performance of academic tasks is always being exchanged for some form of currency—we may call that a grade. This is the symbolic commodification of intellectual labor. In the US, high school students who aspire to attain higher education will submit college applications that reduce them to signs among other signs. The unwavering pursuit of advanced placement courses, internships, participation in student government, all in the name of academic excellence, temporally disjoints students from the present as they become libidinally invested in liberal futurities. Thus, students are barred from becoming more of themselves, more than the potentiality of their intellectual labor. The carcerality of cybertime bars students from moving beyond the hailed sign of productivity. Neoliberal regulation of public education writ large has unfortunately created simulations of educational experiences. In the absence of funding and active investments of care from the state, education has become standardized, a game of memorization. Paulo Friere echoes the notion of a "banking model" upon which contemporary standards for education stand—how much information can be deposited into the minds of students? Unsurprisingly, more students today are painfully dissatisfied with educational experiences, specifically the university; it is all work and no play. Their intellectual labor validates the exchangeability of their sign value in relation to others by overidentifying with productivity, avoiding the possibility of failure, and achieving perfect grades at all costs. Such attitudes are detrimental to the mental well-being of students and only carry over into their collegiate careers. Consider the rise in trends such as sleeping in college

libraries, colloquially known as ‘pulling all-nighters’, surrounded by energy drinks to compensate for the lack of energy required to push cognitive labor forward. Ridiculous as it may sound, this is a typically recognized norm by college students. A more frightening trend, however, is the abuse of prescribed amphetamines (Adderall) to treat attentive and inattentive deficit disorder (A.D.D.). Upon reaching affective exhaustion, the point of cognitive meltdown, one might seek the aid of performance enhancements to sustain and advance the productivity of the mind. However, Berardi reminds us that even the use of psychoactive uppers has its limits, and one will eventually break down. Not even the yuppies of Wall Street in the 80s could keep up with the rhythm of capital as they eventually crashed.

Semiotic commodification of students occurs at various points throughout the academic journey. Perhaps one crucial point often dismissed as an accepted convention is the dreaded university application. Beyond the listing of one’s academic performance over a period of time is the personal statement. Personal statements are disguised as a mechanism by which some nebulous committee is better positioned to get to know the ‘real you’. Already, there is an error here. The real student in the application is a simulated projection, the hyperreal representation of the student. In one sense, the applicant being reviewed is the true representation - an exchangeable sign, data to be processed in the accumulation of value that directly generates financial gains. In another sense, the applicant is but a data double, the xerox of a potentially real student. Much like Twitter, which restricts the number of characters used to share a post, personal statements are designed to retrieve the gist of a student and a very particular reading of a student. Personal statements typically prompt students to write about struggle, hardship, a visceral recalling of suffering, and how they overcame it. This particular experience is highly sought after by application and admission committees of universities. This is not a benign process but the commodification of suffering that feeds symbolic economies of pain. That is, the reduction of meaningful pain narratives to sign commodities—the symbolic exchange of one’s pain and suffering for admission. The implication here is that validation, and thus recognition as a prospective student at a particular campus, is contingent on the exchangeability of one’s suffering. If one does not have the makings of the ideal student, perhaps their pain will be sufficient to gain admission. And this is not unique to university applications, but the university itself. This is partly what Giyatri Spivak teaches us in her article *Can the Subaltern Speak?*

Returning to Tristam Adam's analysis of semiocapitalism's regulation of the voice, and effects on speech, I'd like to contextualize their work to speech within academic spaces. Recall that Adams' argument is that accelerated speech patterns indicate being immersed within capitalist environments, specifically speech attuning itself to the rhythm of semiotic production. There are limitations to human delivery of speech. I am not referring to novel deliveries of accelerated speech, but the everyday performance of speech acts that attempt to capture the tempos of acceleration. Speech is not solely dependent on a vocal apparatus that enables the voice to intelligibly express syllabic, chronological, and intentional speech. Speech also depends on the mind's capacity to meaningfully organize language into an articulate arrangement. I have already established that the cognitive worker is already undergoing a process of neuroplasticity so that the mind can withstand daily cycles of informational terror. Adams raises a worthwhile question: can the voice suffer under conditions of immaterial labor? Yes, and he identifies "vocal fry" not just as the location but as the enunciation of trauma informed by semiocapitalism. Adams brilliantly sums it up:

> ...the vocal fry is proposed here as a register of the human speed limit being reached. The vocal fry is the sounding out of the disjunct between the finite and human and the exploding and accelerating infosphere of semiocapitalism: human buffering, suffering in an excess of speech speeds. As semiotic and linguistic excess overwhelms the human speech rate cannot be maintained and voice sounds. (Adams 2018, p. 40)

Consider vocal fry as a temporal lag in the delivery of human speech. Adams suggests vocal fry can announce itself as verbalized "ums" or significant pauses between connective signs that are struggling to fully realize themselves in speech. It is not enough to recognize vocal fry as the human voice struggling to maintain acceleration that mirrors semiotic flows, but the violent operationalization of the voice in service of accelerating discursive communication that conforms to the rhythm of semiotic production. And where is this more common than the classroom? In educational settings like the classroom, it is not uncommon for students to hesitate in participating in group discussions due to a fear of 'sounding' incompetent or lacking knowledge in contradistinction to their peers. The classroom is not just a site for potential knowledge production, but competition. These two are at constant ends. The quicker one can produce knowledge,

articulate, and circulate that as semiotic currency within the bodily domain of speech, the closer one is in proximity to signs of "competency," or academic excellence, as these are hailed within neoliberal regimes of schooling—self-optimization over mastering academic content. The mind attempts to accelerate processes of understanding in order to respond to a given prompt. There is a desire for one's voice to be expressed, but not enough time in the attention economy to adequately transfer one's thought through the vocal apparatus. As such, students will often go "um," "uh," or simply pause, demonstrating the voice actively coping with the hyper-accelerated rhythms of semiotic sound production. It is the body demonstrating its "vocal limits." In a classroom, to speak is to make relevant contributions such that they keep the rest of the class on schedule, so that it remains productive. Thus, if one wants to speak, one must be able to stay on beat.

BAUDRILLARD PROPER

Why Baudrillard? Of all his contemporaries who composed the mighty intelligentsia of 1960s France, Baudrillard is regarded as an unserious academic who commits intellectual irresponsibility (i.e. *The Gulf War Did Not Take Place*) or a fatal obscurantist who fails to seriously challenge the system of capitalism. These are debatable positions, but I am not interested in such exercises. Instead, suffice it to say that Baudrillard's efforts in formalizing a semiological system that accounts for the semiotic currents of capitalism, combined with his theory of simulation and elaboration on thanatophobia—he is perhaps a proto-theorist of semiocapitalism. Berardi recognizes this in Baudrillard's corpus when he writes:

> There is no more truth, only an exchange of signs. Only a deterritorialization of meaning. In *Symbolic Exchange and Death*, Jean Baudrillard says that the whole system is falling into indeterminacy. This is the essential shift from industrial capitalism to semiocapitalism: indeterminacy takes the place of the fixed relation between labor-time and value so that the whole regime of exchange falls into an aleatory system of floating values. (Berardi 2012, pp. 85-86)

Baudrillard predicted the end of production, or the new schema of endless hyperreal production—not of consumable goods, but signs at a symbolic level. Accelerated simulation *is* the overproduction of meaning that is eventually subsumed by symbolic commodification. As a

student and friend of Felix Guattari and Gilles Deleuze, it may seem strange to include Baudrillard in conversation here. My overall aim is not to reconcile Baudrillard with Deleuze and Guatarri. Rather, for this article, I am more interested in considering Baudrillard as an integral thinker in expanding an understanding of semiocapitalism and how to combat it.

DEATH OF THE REAL

"Welcome to the desert of the real," Baudrillard announced in *Simulacra and Simulation*. First, there was the territory of the real, then its model, or copy. The withering away of the territory ultimately survived by the model means that the model stands in for the real—creating a simulation process whereby the real experience is facilitated by pure simulacrum. It is the disappearance of Ozymandias' empire; all that remains is the desert of the real and mirages that present themselves as genuine articles. Baudrillard does not provide a specific time in which the real has disappeared, but if there ever was one, the overproduction of simulation is as real as it gets. Thus, mapping the coordinates of possible knowledge and experience through appeals to intellectual transcendence is absurd. That is to say, there is no greater truth outside of simulation, no escape hatch that leads one to a higher plane of reality, only deeper levels of simulation. Indeed, as Baudrillard puts it, "it is all of metaphysics that is lost" (Baudrillard 1994, p. 1).

Simulation enables the emergence of hyperreality, or the intensification of the real—it is more exciting, pleasurable, true, and appealing in other respects. It is the "precession of simulacra" made possible by advanced telecommunications that mediate the social, becoming a crucial feature of postmodernity. For example, it can be said that Trump is America's first postmodern president precisely because he is a hyperreal figure. That is, Trump as a celebrity personality is more desirable than Trump, the man, or Trump as a respectable politician. The public has decided it finds the drama of Trump more appealing than whatever the real nature of politics is—like that of reality TV; spectators too will find it more desirable than the boredom of the real. As such, politics becomes something akin to that of a television drama—snapshots and soundbites of political theater that derails the efficacy of political discourse within communication as the hegemony of the sign subsumes it. Another example can be found in today's instantaneous access to pornography. Baudrillard comments on this, the instantaneity of immediate consumption of pornography as a perversion of sex. That is, sex as an experience of shared intimacy and pleasure is what occurs in the real. However, the virtualization of pornography has frightfully minimized

one's proximity to the act of sex, enabling a hyperreal experience of sex. The representation of sex that pornography offers is so far removed from the real that it not only becomes that much more stimulating but more preferable than the sex of the real. The ideal image of man and woman performing before one's eyes in painful transparency clouds the imagination's capacity to differentiate sex as that private act and illusory simulation carried out by performances of adult actors. This is demonstrated by incel culture that is best characterized as a breeding ground for violence against women (physical, psychic, discursive, etc.), where one's epistemic locus of romantic relationships is grossly informed by the mass consumption of mainstream media that depicts the objectification of women, or the hyperreal presentation of romantic and sexual relationships that saturate the hyperreal (i.e. the image chronologically succeeding the real). It is the pornographic performance that precedes the sex of the real.

SIMULATED EMANCIPATION & THE WAGER OF DEATH

Earlier, I mentioned the impossibility of communication as semiocapitalism creates an overproduction of meaning mediated by financialization. As a result, discursive acts facilitated by digital mediums are already undergoing a process of a-signification—we are communicating, but that communication is vacuous. I'd like to return to this notion that Baudrillard explores in *Simulacra and Simulation*.

I am concerned by trends in political engagement that are distinctly marked by the language of consciousness-raising. Meaning, radical gestures are made at a communicative level mediated by mass media technologies. If the oversaturation of information necessarily effaces meaning, then it also becomes divorced from its semantic efficacy. Baudrillard reminds us that radicality and subversion are always at risk of being co-opted by simulation. This is the irony of consciousness-raising in the age of semiocapitalism—these gestures demonstrate affective investments in financialized modes of communication that stifle radical challenges to the system. Consciousness-raising further facilitates the proliferation of signs whereby the information being disseminated is reducible to its own sign that can be exchanged for another—that of being politically conscious or pleasure gained from a kind of 'feel-good' politics. Consider the recent "non-event" of the self-immolation of US soldier Aaron Bushnell, which the world has witnessed as a radical challenge towards US aid of Israel in carrying out genocidal violence against Palestinians, a mirror of indifferent brutality held up to the West. This is a notion explored in *Symbolic Exchange and*

Death. A disruption, no doubt, but perhaps in vain, as the system has chosen to respond in two ways: 1. That soldier's radical suicide was assimilated into another image circulating the infosphere. His gesture has become another sign (pictorial, discursive) that is freely exchanged within mass media platforms. In fact, it did not take long before news outlets deradicalized the intent of his self-immolation as headlines pushed a different narrative, that of another veteran who struggles with mental wellness and succumbed to personal demons. 2. It did not take long before users of social media platforms circulated his image as a kind of social currency that simulates and validates one's level of radicality. Moreover, a cartoonish artwork soon emerged that sanitized the impact of that soldier's fatal strategy. The greater point here is that a notion of marketable potential has already mystified gesture. It has become divorced from its referent, and now, as a sign-commodity, it has become simulacrum.

Baudrillard cites violent acts, like suicide, as subversive acts that challenge the system on the level of the symbolic. To be clear, he is not openly advocating for people to wage terror or commit acts of suicide. Berardi will cite suicide as a very real response, not to the system, but rather a natural consequence of complete fragmentation of the subject, being unable to cope with constant productivity. Though suicide still holds symbolic power, it is not a strategy Berardi will advocate for as he is concerned with creating communities of care and therapy, as proposed in *After the Future*. What needs to be done, then? What shall become our grammar for resistance? Baudrillard proposes a few interesting - let us call them games - ways to counteract the overproduction of simulation in the form of weaponizing death against the system, seduction, reversibility, and fatal strategies. These deserve greater elaboration that I cannot do justice in this article, so perhaps they are better saved for future analysis. Berardi offers suggestions that are aligned with poetic gestures, counteracting cyber time, and acts of economic "desertion." I will elaborate on the prospect of these suggestions in the following section.

AGAINST PRODUCTIVITY

In *The Uprising,* drawing upon Baudrillard's theory of symbolic exchange, Berardi suggests that our task is to free the sign from processes of financialization. If language is ultimately reducible to signs to be exchanged for other signs, then perhaps there is a possibility of reversing this process such that language can escape commodification. Berardi writes,

> …language exceeds economic exchange. Poetry is the language of *nonexchangeability* [my emphasis]...the return of the sensuous body of language. I'm talking about poetry here as an excess of language, a hidden resource which enables us to shift from one paradigm to another. (Berardi 2012, pp. 139-140)

Berardi is speaking of a poetic recuperation of language such that it resists the trap of productivity. Suppose semiocapitalism subjects cognitive workers to depressive, numb states. In that case, poetry is the disruption of informational flows that act as deliberate traffic stops, allowing cognitive workers to not only counteract the speed of cyber time but reclaim time as belonging to their person. Within the disruption poetry creates against the rhythm necessary to consume and make sense of sign consumption, we may find new rhythms to live, breathe, and become aware of the body. What does it mean to become aware of the body? Semiocapitalism drains energy from cognitive workers, sensibility, and sensitivity itself. Because the mind is dulled by intense cycles of semiotic production and consumption, depression forms and limits the cognitive worker's capacity to experience sensation, and organic excitability feels foreign, as one's time is continuously invested in being plugged into digital networks. As such, it is not surprising that there is less time for one to become familiar with one's body—instead, the body is forgotten, the body as possibility of reactivating the affective registers of life. Berardi continues, "Sensibility is in time, and we need to understand the hypercomplex communication of the body" (Berardi 2012, p. 144). Recuperating time seems like a necessary step in challenging semiocapitalism. It is not solely the emancipation of time for oneself, but shared time between other bodies, creating shared expressions of energy that renew psychic life for the cognitive worker, but also hope for social solidarity—the corporeal emergence of the general intellect. Berardi's understanding of poetry reminds me of Baudrillard's commitment to playful writing styles—that is, appropriating the acceleration of simulation to create modes of communicative exchange that resists definitive interpretation, slipping into unintelligibility.

I am more convinced by a strategy of disidentifying with the sign of productivity. Remaining within the prescription of reclaiming time for oneself, we must promote a politics of therapy as Berardi suggests in *After the Future*. Therapy must be a conscious disinvestment from desiring productive labor, if only through momentary gestures. This will, in some sense, allow a psychosomatic healing process to begin. Berardi uses the word "desertion" to elaborate the

necessity to withdraw from the infosphere. Much like a soldier who deserts their military duties, we must also adopt the same stance to find new modes of existence beyond productive labor. With the recuperation of energy, we will have an easier time becoming self-creative subjects that can realize new horizons, new imaginations, and redefined forms of leisure beyond the obstructive paradigm of semiocapitalism. Embrace lag time, even if suspension of time can only be found in the depths of one's slumber.

Bibliography

Berardi, Franco "Bifo" (2012). *The Uprising: On Poetry and Finance.* Semiotext(E). pp. 19, 85, 86, 139, 140, 144

Baudrillard, J. (1994). *Simulacra and Simulation* (S. Glaser, Trans.). University of Michigan Press. p. 1

Adams, T. V. (2018). "Accelerations and speed limits: An essay on the vocal limits of semiocapitalism," *Journal of Interdisciplinary Voice Studies*, 3(1), 21+. https://link.gale.com/apps/doc/A539923120/AONE?u=anon~ffcace2e&sid=googleScholar&xid=2dab5a8d. pp. 33, 38, 40

HUSSERL, NOEMA, AND THE STRUCTURE OF THE OBJECT

Alexis Valenzuela Haro

INTRODUCTION

Husserl's quest to make the study of experience, phenomenology, a rigorous scientific enterprise pushed him to develop new terminology. The concept in phenomenology introduced to study the objects of our experience is called the noema.

Traditionally, the noema has been conceived as the object of our intentional consciousness *as experienced* (Drummond 1990; Crowell 2006). To say that consciousness is intentional means that at least some mental states are such that they can be characterized as being *about* objects. The noema would then be the object of perception *as* perceived, the object of memory *as* remembered, etc. The object we intend and the noema are the same object given from a different perspective.

On the other hand, the Fregean interpretation (Føllesdall 1969; Smith and McIntyre 1971; Smith 2023) sees the noema as an abstract object that mediates the relation between us and the object we intend. Just as a term's sense bestows linguistic acts with reference, the noema as an intensional entity broadens the concept of sense beyond the sphere of linguistic acts to that of all intentional acts of consciousness.

According to the Fregean interpretation, phenomenology attends to a mental object called the noema that fixes the reference of our intentional acts of consciousness and renders perceptual intentional episodes amenable to linguistic sense. I will argue that this interpretation misconceives what Husserl means by object by failing to differentiate between a representationalist and transcendental conception of an object.

To develop this argument, in §1, I will introduce Husserl's precursor to the noema, act-quality and act-matter, to motivate the introduction of the noema. After introducing the noema, I will explicate its logical structure into 3 components: noematic Sinn (determinable X), noematic core, and the thetic component.

Next, in §2, I will argue that the Fregean interpretation fails because it treats intentionality as an ontological relation between a mind-independent object and a mental representation. Then, I will analyze Kant's Copernican revolution to argue that the notion of object entailed by a

transcendental framework is different from representationalist accounts such that we can map the noema's structure to the transcendental object. In this section, I map the thetic component to the surface level of the transcendental object.

Finally, in §3, I will show how Husserl's concepts of horizon and essence allows him to reverse the Kantian Copernican revolution from within a transcendental framework. This allows him to differentiate between two levels of the object (pre-predicative and predicative) to provide a robust account of the process of objectification; the transition between levels of the object. I will relate the two levels of the object considered transcendentally with the noematic Sinn and noematic core.

§ 1 Act-Quality and Act-Matter

Phenomenological analysis distinguishes between the intended object and the object as intended. The intended object is the object that our mental intentional act relates us to while the object as intended is the object as experienced. Originally, Husserl's mentor, Brentano, had devised a theory of intentionality according to which he claimed that the intended object and the object as intended were one and the same object. However, this led to the consequence that when we think, "I'm going to have drinks with my friend tomorrow," the intended object is not your friend, but the mental representation that is the real psychic content of experience as sense-datum that is part of the experience. The real psychic content is the part of the experience that is given as genuinely experienceable content; for example, an object would not be part of our psychic content because the object is not literally a part of our experience, but according to sense-datum theories, sense-datum is a real part of the experience.

In the Logical Investigations, Husserl tries to overcome this absurdity by providing a new account of the structure of the intentional act. The intentional act was broken into two components: the act-quality and act-matter. As instantiated in real psychic acts, the act-quality is the part that determines the psychic modality (i.e. perception, imagination) in which we intend the object. The act-matter is what determines the object of the act. In contrast to Brentano's object-theory of intentionality; the act is differentiated via act-matter because the object cannot play this role since it is not immanent to our real psychic processes.

Husserl called the act-quality and act-matter the intentional essence as viewed from the intentional content of the act, rather than the real psychic content of the act. The intentional essence

of the act allows for the transmission of ideal objectivities that transcend any real psychic subject. For example, when we say that, "I perceive that apple" or "I imagine that apple" we are intimating the way our psychic processes comport themselves towards the object and expressing an ideal objectivity that can be tokened in different psychic episodes. The object is merely determined by the immanent psychic processes but is no real part of the act as some mental psychological entity.

However, the intentional essence of the act followed a model of ideal instantiation. The ideal content of the act was tokened into particular psychic states. By posing the problem in terms of the instantiation of ideal entities, Husserl was unable to account for the relation between immanent psychic states and ideal transcendent objectivities. Hence, Husserl turned to transcendental idealism both to overcome the problems of Brentano's theory of intentionality and dispense with the ontological formulation of act-matter and act-quality for a transcendental formulation of noema and noesis. Before we get into what this means, I will introduce the noema and noesis, the structure of the noema, and the literature surrounding it.

§ 1.1 The Structure of the Noema

After his turn to transcendental idealism, Husserl conceived of the intentional act as composed of two components: noema and noesis. The noeses contain the real psychic content of an intentional act. This includes the sensory data, hyle, of perception and a functional component that "bestows" the hyle with sense and animates it into a unity of sense. The noematic content, by contrast, includes the intentional content of the act that is the bestowed sense. This sense that belongs to the noematic component determines the meaning (Sinn) of the act whether it be a linguistic or perceptual act.

The noema consists of three logical moments: the noema's thetic component, the noematic Sinn, and the noematic core. The noematic Sinn is the determinable X ('something' that is the bearer of predicates), and the noematic core is the object's unity of predicates as distinguished from the determinable X. The thetic component is the mode of givenness by which we intend the object and the mode of being of the object as given to us.

The Fregean interpretation treats the determinable X as a demonstrative that refers to its intended object directly. The determinable X's role is to transform the core into a demonstrative that refers to its intended object directly as opposed to a reference mediated by a description (one of the core's predicates) (Smith and McIntyre 1982, p. 204). If our intentional object is an apple

on our desk, for example, we might express the sense of the perceptual episode as: “I perceive something which is red and circular and...” The thetic component is the perceptual mode of givenness, the cluster of predicates of the apple is given and reference to the apple is determined according to its predicates. However, the same perceptual episode might be expressed as, “I perceive *that*,” if the determinable X refers to the object directly.

The Fregean interpretation commits itself to the view that the object as intended, and the intended object are not equivalent with one another. The object as intended, the noema, provides us with the intensional sense that renders a linguistic objectivity expressible because much like the act-matter, the noema’s role is in determining the reference of the act. In cases where there is no actual intentional object, there is still a noema that would have directed our act *as if* there had been an actual existent object.

Gurwitsch, the originator of the traditional interpretation, identifies the object as intended and the intended object. For Gurwitsch, the object is a continuum of appearances organized according to a rule (a Kantian Idea). Gurwitsch also appeals to a Husserlian concept, the idea of a horizon. Horizons are the part of the experience that predelineates future possibilities; they implicate other aspects of the object. Through horizonal intentionality, the act can intend other parts of the objects implicated in its horizonal structure. Gurwitsch ends up providing an account of the noema that conceives it as a part-whole relation between the appearances of the perceptual act, the momentary noema, and the whole object, the concrete noema. According to Gurwitsch, we only experience appearances and the appearances’ horizonal implicates.

Drummond agrees with Gurwitsch that the noema is the same as the intended object, however, he thinks that Gurwitsch’s account fails because appearances and objects are incommensurable. For Gurwitsch, the intended object cannot be given in the act because it is an infinite limit that regulates the system of appearances. Drummond will differ from Gurwitsch because he introduces identity-in-manifold analysis in contradistinction to whole-part analysis. The difference is that whole-part analysis is static, while identity-in-manifold analysis allows Drummond to bring in the notion of temporally extended intentional acts. In this way, he takes the determinable X as characterized in teleological terms:

> The determinability of the "X" should be understood as the object's ability to come to a more precise determination in the course of a temporally extended experience (cf., e.g.,

DR, §§27ff., *APS,* 5, 20-22); as such, the "X" of *Ideen* I is a genetic concept illegitimately confined within in a purely static account [1] (Drummond 1990, p. 155).

Drummond envisions the X as an object that is fully given in experience and implicates further predicative determinations in the mode of its thetic givenness. The object exists through the temporal manifold of appearances, not over them as an unfulfillable goal, but as given "in" through the appearances. Therefore, Drummond ends up rejecting the linguistic demonstrative account of the noema found in Smith and McIntyre.

I will defend the conception of the noema present in Drummond. I will take a Kantian approach in doing so. My strategy will be to argue that the Fregean conception of the noema uses a representational account of an object, while the traditional approach remains faithful to Husserl by using a transcendental account of an object. I will unpack the structure of the transcendental object through an analysis of the Kantian Copernican revolution and map the structure of the object of transcendental philosophy onto the noema.

§ 2 Representationalism

Representationalism construes "mind-world relations as relations between internal representations to external object" and introduces a gap between mind and world that is supposed to be crossed by postulating emergent normative properties intrinsic to the brain (Jansen 2014, p. 79).[2]

I take the Fregean interpretation of the noema to be decidedly representationalist insofar as it (1) is a mental entity, (2) treats non-existent objects as "bad" correlates to consciousness, and (3) has normative properties intrinsic to it.

(1) is a result of the denial that the noema is the intended object. The only place left for the noema is a mediating mental entity between us and the object we intend. (2) speaks to some of the underlying commitments of the Fregeans. They think that normative properties are needed to establish the reference with the object because some objects make for "bad" correlates to consciousness. By this, I mean that the Fregeans prioritize spatio-temporal particulars as "good" correlates of consciousness. To say that there are intentional acts fulfilled by object-less noema is to say that there is a gap between mind and world because sometimes mind is enough to fulfill itself. (3) requires that we preload the mind with the subjective structures needed to establish a reference to the object.

The Fregean interpretation commits itself to a semantic internalism that sees the noema's intensional nature as providing the conditions for reference (Crowell 2006, p. 338). Internalism is the view that reference is determined by narrow internal (psychological) content. According to Crowell, insofar as internalism is representational, it is committed to the view that representations are 'ontological' entities. The notion of object in representationalist accounts is only accidentally related to the mind because representationalism treats mind and world as two separate categories mediated by a third term, like a representation.

In § 2.1, I will analyze the Copernican revolution in transcendental philosophy to distinguish two senses of the object; the object operative in the representationalist framework of Realism-Idealism and the new sense of object entailed by the transcendental framework.

§ 2.1 Copernican Revolution

Transcendental idealism positions itself in opposition to the Realist-Idealist framework. The Realism-Idealism framework is, "…a traditional dualist framework in which items are exhaustively either empirical mind-dependent ideas or empirical mind-independent external objects" (Bird 2006, p. 288).[3]

The relation between subject and object is that of epistemic grounding where our knowledge of one term is grounded *in virtue* of our primitive insight into the other term. For example, in realism, we acquire knowledge of the subject based on a primitive concept we have of an object independent of the subject. Gardener describes the primitive sense of the object as the concept of an object in general, and having assumed such a concept, in what way we stand in an epistemic relation to it (Gardener 1999, p. 25).

The relation requires a mediating term that guarantees a link between the subject and an object because they are both different *kinds* of things. The subject's mind-dependent ideas are inner mental representations that relate us to mind-independent objects that are external and physical. I take the Fregean interpretation to have an implicit notion of the primacy of the external physical object, so the noema must be third term (the normative conditions of the representation) in a triadic theory of intentionality to overcome the ontological gap.

By contrast, transcendental philosophy does not treat any notion of object as primitive because it provides a principle that precedes and grounds the subject and object relationship. The sense of the relationship of subject to object is determined according to an a priori principle, and

thus, object is defined functionally in relation to it's a priori and essential relation to the subject. I will discuss the Copernican revolution to show how we can dispense the need for a middle term if the relation between subject and object is related necessarily according to an a priori principle.

Just as the relation between the sun and the earth required the law of gravitation to functionally define each term in relation to each other, so Kant's Copernican revolution requires a more general principle that precedes and contextualizes the relation between subject and object *without* treating either term as epistemically primitive. Kant's principle, the transcendental unity of apperception (TUA), conditions the relation of subject to object as a necessary relation a priori. There is no need for a mediating term to link the subject and the object because the possible conditions of our experience contextualize any possible relation we can have with objects.

The TUA states that any objectively valid representation of objects must be such that representations of experience are linked together necessarily by the categories, the pure concepts that synthesize our sensible intuitions and make experience intelligible in unified self-consciousness. Synthesis is the activity that links together various presentations given to us passively by our sensibility into a representation conceptually structured according to the categories.

The unity of the representation is the condition that allows the representation to be ascribed to a numerically identical subject because the synthesis that unifies the object of experience is the same synthesis that unites self-consciousness. Thus, " … whenever we represent something as an object or objective event, self-consciousness is at least involved as an awareness of our so representing" (Schulting 2021, p. 3).

The a priori principle, TUA, determines the sense of object only insofar as the object is mutually related to the unity of self-consciousness. TUA grounds the reciprocally determined sense of subject (self-consciousness) and object such that self-knowledge is at the same time knowledge of the object. The a priori rules that provide determinacy to the object are the same a priori rules that unite self-consciousness; thus, the categories that determine the object can be deduced from the unity of self-consciousness. As Kant says, "The 'I think' must be able to accompany all my representations" because subject and object share the same grounding principle (1996, p. 177).

Thus, we may distinguish two notions of object. Object$_1$ is the precritical notion of object, the object that is in-itself independent of our experience. This notion of object is only accidentally

related to the subject because it frames the mind and world gap as ontological and requires a mediating term to connect them. Object$_2$ is the postcritical notion of object, where the object is immanently related to the unity of self-consciousness according to an a priori determination. It does not need mediating entities to bridge an ontological gap because it is directly correlated to our experience.[4]

§ 2.2 Transcendental-Phenomenological Reduction

In this section, I detail Husserl's transcendental-phenomenological reduction to show how Husserl enters the domain of the transcendental and switches from object$_1$ to object$_2$. Husserl thematizes transcendental consciousness with a methodological technique he calls the transcendental-phenomenological reduction. The reduction provides phenomenologists with a way to acquaint themselves with the noema. First, we bracket the general thesis of the natural attitude. In the natural attitude, we posit the thesis that the world exists ontologically independent of our subjectivity as an external "out there" against an internal insulated subjectivity that is "in here."

In the reduction, phenomenologists perform the epoché; the suspension of all judgments concerning the existence of all transcendences. We bracket the judgment of existence to examine the appearance conditions of the phenomena that we take as naively existent. We don't doubt the act; rather, we cease to live through the act and view it from a vantage point disconnected from it. We bring the noema to our attention by thematizing the immanent profiles of the object (whether it be a spatial object with multiple sides or a temporal/ideal object with horizons that escape the possibility of attaining the entirety of the object with a single intentional predicative ascription).

Husserl rejects the inner and outer distinction that characterizes representationalism's hidden presupposition of a world of objects against us. We dogmatically accept object$_1$ as a primitive existent that is given with sense prior to subjectivity. We shift our attitude from treating the object as an object$_1$ to an object$_2$. It is an object insofar as it is genuinely transcendent to us; it is not "in" our mind because its sense cannot be obtained in one single psychic intentional act. A form of phenomenological immanence is distinguished from the psychic immanence that plagued Brentano's account.

With the rejection of the natural attitude, we no longer view consciousness as literally in the head, but instead as the very opening whose activity reveals the constitution of objects. Thus, we can return to the identity of the object as intended and the intended object without the pitfall of

Brentano's theory of intentionality. The object as intended is not a real immanent psychic constituent of our experience; rather, it is an object phenomenologically immanent to a consciousness that is more akin to a place that embraces the world.[5] This picture of consciousness emerges in the rejection of the inner and outer dichotomy that characterizes mind and world as two separate things.

Thus, the first part of the transcendental object corresponds to the thetic component of the noema. It is in the essence of the transcendental object that it is essentially correlated to experience. The noema emphasizes this correlation by thematizing the modalities (i.e., perception, imagination, etc.) in which the object is given as being part of the noema itself.[6]

The object's relation to the unity of self-consciousness and the a priori determination of the object will allow Husserl to distinguish two levels of the transcendental object that will differentiate transcendental phenomenology from Kantian transcendental idealism. In the next section, I will explicate the levels of the object and relate them to the determinable X and the noematic core. The a priori principle and the levels of the object will allow us to locate the norms of the intentional act on the object, i.e., noematic side of the experience.

§ 3 Horizon and Essence

For Husserl, the a priori principle that motivates synthesis will be found on the side of object$_2$, instead of a priori rules located in our mental faculties. The rules will be available for phenomenological introspection in the object as a "node of intentional implications" (Crowell 2006, p. 344). The object possesses horizons that entail other features of the object's structure. For example, an apple may 'entail' its material backside, its formal categorial properties such as unity, plurality, relation, etc. For Husserl, the world is itself structured categorically; it is not a relation that is imposed by the subjective conditions of the ego.[7]

Unlike Kant, the ego's synthetic activity only reveals the object and lets it self-constitute. As Smith explains, "If we say the egret "is constituted" in my experience, one may want to ask: what does the constituting? Is it I qua subject, or is it the act qua intending? No, nothing like that. The literal rendering is "constitutes-itself," or "self-constitutes," but there is no entity that does the constituting" (Smith 2023, p. 71). Self-constitution is the explication of a pre-given meaningful object; in articulating the object's structure, we are not determining it according to mental categories but discerning the object's own structural components.

I call the thesis that a subject's synthetic activity is determined by the demands of the object$_2$ the Copernican Reversal thesis (CR). In representationalism, the subject conforms to object$_1$ according to a correspondence between object$_1$ and the representation; in CR, synthetic activity conforms to the demands of object$_2$. In other words, the norms that determine the object and its material and categorial structure are located a priori in the world.

de Warren explains Husserl's Copernican reversal:

> It is therefore not consciousness that unifies the manifold of givens into an object, but the object that guides the manifold of "non-identical consciousness"- a consciousness that only itself gains unity by virtue of its intentional object. In this manner, Husserl quietly reverses Kant's celebrated Copernican Revolution without thereby returning to any kind of pre-Kantian position (de Warren 2015, p. 243).[8]

Kant cashes out his a priori principle in terms of a conceptual scheme that relates self-consciousness and its object, while Husserl cashes out the a priori principle in terms of the pre-givenness of objects that unite a multiplicity of intentional acts; it is the object that unites self-consciousness.

The invariant part of the object that provides the stable self-identical pole around which intentional acts unite is the essence of the object grasped by virtue of horizons. The object's horizons guide our perception around the object according to a process where we anticipate future profiles of the object and fulfill them when our expectations are met.

Horizons "allow essences to be grasped, but the notions of horizon and essence diverge" (Zhok 2012, p. 126). Horizons relate to the perceptual structure of experience, being essentially in the form of anticipation and fulfillment around a set of possibilities. In contrast, essences are "ontological thresholds, primordially rooted in our motivated confrontation with sensuous transcendence" (Zhok 2012, p. 99). The essence is part of the object's invariant structure that unites empirical apperception by providing the constancy necessary to relate the ego to itself at different points of time.[9]

The object essence that unites empirical apperception is the same a priori essence that bears transcendent intentional implicates that can be brought to fulfillment. Essences reveal the co-essential nature of consciousness and transcendence because the same essence that unites self-

consciousness is the essence that implicates different determinations of the same object under a specific form of thetic positing. The centaur, as a non-existent object, implicates further determinations in the manner appropriate to its mode of givenness, the imagination, and unites self-consciousness *because* it implicates further determinations that attempt to fulfill an anticipation.[10]

§ 3.1 Object Levels

In this section, I distinguish between two levels of object$_2$ to diffuse the Fregean motivation for introducing normative conditions to a mental representation. The object is not impressed on us as a ready-made objectivity structured predicatively, but we are given the object passively in a field of sense, Husserl tells us, " ... It is not mere particular objects, isolated themselves, which are thus pregiven, but always a field of pregiveness, from which a particular stands out and, so to speak, "excites us" to perception and perceptive contemplation" (Husserl 1939, p. 72).

In Kant, the bare given is nonconceptual content differentiated solely according to a priori spatio-temporal individuation conditions.[11] Without the categories, sensible intuitions would be blind and not fit for discursive knowledge. However, in Husserl, the field of pregiveness is meaningfully differentiated prior to the conceptual subsumption of particulars in judgments.

The fields of sense are constituted in passive syntheses prior to the active syntheses that constitute the predicative object. Passive syntheses are syntheses that have the lowest level of activity and prepare our encounter with objects; they precede ego involvement and condition the objects that call the attention of the ego. Objects in the field affect us by possessing a significance that motivates the ego to explicate the object contra sense-datum theories that endow objects with a causal power to impress themselves upon us.

I would now like to differentiate between two levels of object$_2$. Object$_{2a}$ is the object that lies vaguely determined in the field of sense. The field conditions it as significant prior to predicative explication. Object$_{2a}$ is not yet propositionally articulated with determinate inference rules; rather, it opens an interpretative space where the object's power of affection is based on a distinction between familiarity and unfamiliarity. At this preconceptual level, we navigate the environment according to a sense of familiarity that attracts proper kinesthetic action (van Mazijk 2016, p. 429). Object$_{2b}$ is the object that has been explicated and consolidated into a bundle of

predicates. The more familiar the object is to us, the more it can draw in the ego to determine its predicative structure.

In a strict sense, object$_{2a}$ is not an object at all, as it has not been constituted into a linguistically repeatable objectivity. However, to emphasize the continuity between the motivating source of the object and the fully fleshed out object, I have called it an object because we can recognize it as an object after the fact.

Husserl himself seems to sometimes call object$_{2a}$ an object prior to its linguistic repeatability and indeed require that it be pregiven before a judgment is made: "Every judging presupposes that an object is on hand, that it is already given to us, and is that about which the statement is made" (1939, p. 14). In contrast, the pregiven for Kant appears to be incommensurable with the conceptual object as determined by the categories, however, Husserl's theory of pre-predicative experience seeks to make the pregiven object commensurable with the object determined by judgments such that we can identify the pregiven itself as an objectivity prior to any determination after we have clarified the object in judgment.

For example, Husserl says, "We can also say that before every movement of cognition the object of cognition is already present as a dynamis which is to turn into an entelechia" (1939, p. 29). For Plato and Aristotle, dynamis was the concept that was supposed to unite the notion of incommensurable matter with the actualized object. Dynamis was the potential activity present in the matter that was otherwise incommensurable with the form that tamed it. It could be actualized by a form into an object (enetelechia) only on the assumption that matter possessed the characteristic of potential activity, dynamis (Chiurazzi 2021, p. 39).

By characterizing matter as potential, Aristotle can link two unlike things together. In the same vein, I'm arguing that Husserl is using the idea of passive synthesis to make the object in the field of pregiveness commensurable with its determined form in judgment without resorting to a formal principle like the TUA. The object in the field of pregivenness can be thought of as an object commensurable with the objectivity proper to the ideal repeatable linguistic objectivity *because* the field of sense is a field of unactualized potentials (dynamis).

In a roundabout way, the noema's forerunner, act-matter, is aptly named. As Drummond said, we may characterize the determinable X teleologically because it has the potential to come to more precise determinations just like the dynamis proper to teleological characterization of Aristotelian matter.

To make the point clearer, we can compare Husserl's treatment of the issue with Kant's. For Kant, prior to the synthesis responsible for self-consciousness, our senses provide us with an indeterminate manifold. Objectivity is a function of self-consciousness and nothing without self-consciousness. The objectivity of object$_2$, according to its categorial determinations for Kant, is incommensurable with appearances. This objectivity doesn't retroactively become applied to appearances. However, self-consciousness as a multiplicity of noetic acts unified by a noema provides an objectivity commensurable with appearances insofar as consciousness doesn't provide the objectivity to appearances, but the synthesizing act assents to the teleologically characterized determinable X and aims to unify the object's different predicative determinations in a temporally extended experience. The objectivities that belong to appearances are revealed by the passive syntheses that render the field of sense determinate via comparison, contrast, and association.[12]

This distinction between pregiven object and determined object corresponds to two levels of "contemplative perception" that Husserl differentiates. The first level, "simple apprehension and contemplation is the lowest level of common, objectifying activity … " where object$_{2a}$ is posited as existing according to the passive syntheses that structure the field of sense (Husserl 1939, p. 104). Object$_{2a}$ is something that lies in our perceptual field vaguely, but our attention has not fully turned to it. It draws us because it *stands out significantly*, rather than causally. Thus, object$_{2a}$ is the transcendental correlate of the determinable X. Transcendentally, it serves as the principle that unites empirical apperception, and as determinable X, it is the motivating source of the object that will become clarified in a temporally extended experience.

The second level, explicative contemplation, "strives to explain all that it "is," what it manifests of itself as regards to internal determinations, to enter into its content, to grasp it in its parts and moments, and to enter anew into these by taking them separately and letting them display themselves…" (Husserl 1939, p. 104). An object, an apple, is grasped actively as object$_{2b}$, rather than passively, such that we can grasp its parts in separate judgments, "the apple is red" and "the apple is round." Object$_{2b}$ serves as the transcendental correlate of the noematic core. Transcendentally, it serves to indicate the possible transitions of self-consciousness according to horizonally intended predicative implicates, while as noematic core, it provides the a priori essence that regulates all possible predicative ascriptions of the object.

We correctly perform the process of objectification, the transition from object$_{2a}$ to object$_{2b}$, when we synthesize an object that contains a bundle of predicates that are truly its own. The

process of objectification, and thus the perceptual activity of anticipation and fulfillment, is normative insofar as the self-constitution of an object is a task that we can fail according to norms that belong to the object in its significance to affect us and propel us into explicative action, not norms that belong to a mental representation nor mind conceived of ontologically.

An example of this explicative process is the transition from image to physical object in a perceptual positing.[13] The perceptual noema presents us with an image that becomes constituted as a material phantom in a synthetic coincidence of perspectives of the image. We synthesize the image due to our kinesthetic ability. Husserl calls the phantom the sensible apparition considered apart from its causal and functional properties. Drummond gives the example of an infant who sees a hammer for the first time. The infant does not know what hammers can do, so they perceive an object without any causal or functional properties (Drummond 1990, p. 157). In a temporally extended experience, the hammer becomes constituted as the physical material object as determined by the norms of the object, viewed as the perceptual noema that unites a multiplicity of noetic acts. (Self-consciousness that thematizes the structure of perceptual activity.)[14]

In sum, the a priori essence of the object, grasped in virtue of the horizons of perception, ground the relation between the empirical apperception and $object_2$ by providing the essence that simultaneously unifies consciousness in trying to fulfill an anticipation and entails other parts of the object. The norms that belong to the intentional act are not part of the noema as a representation but are rather part of the object's essence. The object's essence is transcendent to consciousness, but the essence is related to consciousness in an appropriate way because the object (1) unites self-consciousness, and (2) provides the norms that guide the intentional activity of the anticipation and fulfillment of intentional implicates that determine the object. For example, the constancy of the apple given perceptually (1) unites self-consciousness in a way that (2) implicates certain material and formal predicates.

CONCLUSION

In this paper, I have argued that the Fregean interpretation fails as an exegesis of Husserl because it implicitly treats the noema as a representationalist object, while Husserl conceives of the noema as a transcendental object. First, I described the structure of the noema's three components: the determinable X (noematic Sinn), the noematic core, and the thetic component. I related the noema's components to Husserl's early account of the intentional act in the form of act-quality

and act-matter. Next, I explicated the structure of the transcendental object. I argued that the Copernican revolution unveiled a notion of object that was entirely immanent and correlative to experience. Furthermore, two levels of the transcendental object could be distinguished due to Husserl's transcendental Copernican reversal. I argued that we could map the structure of the transcendental object onto the noema by showing how the transcendental object relates an a priori determination and self-consciousness in a way that was structurally equivalent to the logical moments of the noema.

The noema is immanent because the reduction brings us to the sphere of pure immanence and forces us to interpret object_1 as object_2 because it is phenomenologically immanent and essentially correlative with experience (the thetic component). It is a pre-objectified surplus of meaning because we distinguish the determinable X (object_{2a}) from the core (object_{2b}); the determinable X is the motivating source of a judgment that bears horizons that act as norms that overflow the current profile of sense in lawfully determined transitions. The very structure of perception is to bring the absences indicated by the horizons to presence, thereby objectifying the noematic Sinn according to an essence that both structures the unity of experience and the horizonal intentional implicates that belong to the object's structure.

NOTES

1. Likewise, I agree that Husserl's concept of noema entailed that Husserl's thought would move in a genetic direction. Hence, my most cited work in this paper is *Experience and Judgment*.
2. Representationalism is used narrowly in this paper to identify a family of phenomenalist and sense-datum theories.
3. Kant makes the point that we need a third mediating term in the Paralogism of the A-edition, where he argues that when we treat matter as a thing-in-itself, "a third being must intercede in order to bring about between the two, if not interaction, then at least correspondence and harmony" (p. 415, 1929). We distinguish two notions of thing-in-itself, the thing-in-itself ordinarily is the empiricist assumption that matter signifies a thing-in-itself; by contrast the noumenon is the thing-in-itself in relation to eidetic objects. In my argument I use the first sense of thing-in-itself, later in the paper Husserl will come to legitimize the existence of eidetic objectivities.

4. Jansen sees the agreement between Husserl and Kant stemming from both trying to overcome the representationalist tendencies of their contemporaries. For Kant, this was the realist-idealist framework, and for Husserl, this was the empiricism represented by Brentano (2015, p. 83). The notion of object$_2$ arises as a direct polemic against base representationalist assumptions.
5. For a view that argues that this entails the spatialization of consciousness see Shim (2017).
6. What does it mean to say that perception is part of the noema itself? It is intended as a metaphysically neutral statement that the object's mode of givenness is such that it can be given in the mode of perception. We remain neutral between physiological questions concerning the eye and perception; however, the doors remain open for us to adopt an account in which perception is the functionally simple cause of the eye. Such a Bergsonian inspired account is developed in Posteraro (2022), while we remain neutral in phenomenological analyses, the perceptual essentiality of the object can open interesting metaphysical hypothesis of the relation of body (instead of mind) to the world.
7. Rogover (2022) compares the material a priori given in experience to the Kantian formal a priori. For Husserl, eidetic idealities are given genuinely in experience, while for Kant the a priori relativizes all objects to a human conceptual scheme. Rogover states that it is due to the empiricist assumption Kant inherits that experience is an unorganized manifold organized by the subject. My additional claim is that the essence as the material a priori contextualizes the relation between noetic activity and noematic content such that the norms of perception, imagination, etc. can be explicated through an analysis of the possibilities that belong to the givenness of the objects and their material a priori. This is the upshot of my Kantian reading of Husserl, self-knowledge as knowledge of our modalities of thetic positing can grant us knowledge about the structure of the object we have posited. An insight into the norms of perception *is* an insight into the structure of the material object. For this reason, Husserl tells us that not even God could intuit the material object without regard for its adumbrations (2014, p.136).
8. The Copernican reversal thesis may be stated thus, "Apriority is not linked to subjectivity, and still less to its so-called origin in it, but rather corresponds to that which can be given in an intuition of essence" (Schnell 2021, p. 398). For further discussion in the literature see Pradelle (2012); Rogove (2021); Schnell (2021).

9. Lampert (1995, p. 63) sees that the result of Husserl's 2nd investigation is to show that any particular object identification requires the simultaneous identification of the universal. He argues that universals are legitimate objects of experience because individual objects cannot be synthesized without a universal objectification. Thus, I locate the invariant in the universal essence that is required for any particular constitution.
10. The centaur is then not a non-existent object but a fictional object with genuine fictional existence. We may consider Husserl a sort of permissivist about existence for who every object exists genuinely. For a discussion of permissivism see Schaffer (2009). For support of Husserl's permissivism see Husserl's discussion of the principle of principles (2014, p. 92).
11. In this short space I cannot hope to bear on the debate between conceptualism and nonconceptualism in Kant, I only adopt the popular rendition of Kant.
12. Husserl's genetic account of passive synthesis lands him in line with the post-Kantians following Maimon. For Maimon, "The differential elements of the objects are the noumena; the objects that spring from the phenomena" (Atlas 1969, p.110). He sees determinate objects as integrals of the infinitesimal elements of sense perception; the noumenon for Maimon is not an unknowable entity, but instead a Kantian Idea that constitutes empirical synthesis genetically. Prior to the object, the structure of its field must be determined, Maimon thinks, to answer the Kantian question quid juris. (What right do we have to apply a priori categories to experience?) Similarly, Husserl's notion of the determinable X takes the field of sense as containing teleological potentialities that lead to the genetic constitution of the object. Both Husserl and Maimon see that transcendental philosophy must become genetic hence we should treat the noema genetically.
13. In this paper, we have for the most part analyzed the perceptual noema for the sake of simplicity, but other noemas have their own sui generis structure that also must be analyzed on their own mode of givenness.
14. The first part of the image-analysis of this paragraph parallels Gurwitsch's part-whole noematic analysis, and the second part parallels Drummond's identity-in-manifold temporal analysis. There is no contradiction as Drummond's analysis is an extension, rather than a correction of Gurwitsch's.

BIBLIOGRAPHY

Atlas, Samuel. (1964). *From Critical to Speculative Idealism: The Philosophy of Solomon Maimon*. (The Hague: Martinus Nijhof).

Bird, Graham. (2006) *The Revolutionary Kant: A Commentary on the Critique of Pure Reason*. (Open Court)

Chiurazzi, Gaetano. (2021) *Dynamis: Ontology of the Incommensurable.* (Springer Verlag)

Crowell, Steven. (2008) "Phenomenological immanence, normativity, and semantic externalism," *Synthese, 160*(3): 335-354

de Warren, Nicolas. (2015) "Concepts without pedigree: The noema and neutrality modification: Section III, chapter 4, On the problems of noetic-noematic structures," in A. Staiti (Ed.), *Commentary on Husserl's "Ideas I"* (pp. 225-256). (De Gruyter)

Drummond, John J. (1990) *Husserlian Intentionality and Non-foundational Realism: Noema and Object* (Springer)

Føllesdal, Dagfinn. (1969) "Husserl's Notion of Noema." *Journal of Philosophy, 66*(20): 680-687

Gardner, Sebastian. (1999) *Routledge Philosophy Guidebook to Kant and the Critique of Pure Reason* (New York: Routledge)

Husserl, Edmund. (1973) *Experience and Judgment: Investigations in a Genealogy of Logic* (L. Landgrebe, Ed.). (London: Routledge and Kegan Paul)

Husserl, Edmund. (2014) *Ideas for a Pure Phenomenology and Phenomenological Philosophy: First Book: General Introduction to Pure Phenomenology*. (Indianapolis/Cambridge: Hackett Publishing Company)

Julia, Jansen. (2014) "Taking a Transcendental Stance: Anti-Representationalism and Direct Realism in Kant and Husserl" in F. Fabbianelli & S. Luft (Eds.), *Husserl und die klassische deutsche Philosophie, 212* (pp. 79-92). (Switzerland: Springer; Dordrecht)

Kant, Immanuel. (1929) *Critique of Pure Reason (1781-1787)* (K. Smith, Ed.). (London: MacMillan)

Lampert, Jay. (1995). Synthesis and Backward Reference in Husserl's Logical Investigations. *Phaenomenologica, 131*. (Springer).

Pradelle, Dominique. (2012). Par-delà la révolution copernicienne: Sujet transcendantal et facultés chez Kant et Husserl. (Paris: Presses Universitaires de France).

Posteraro, Tano. (2022). *Bergson's Philosophy of Biology: Virtual and the Vital.* (Edinburgh University Press.)

Rogove, John. (2020). The Phenomenological a priori as Husserlian Solution to the Problem of Kant's "Transcendental Psychologism". In I. Apostolescu & C. Serban (Eds.), *Husserl, Kant and Transcendental Phenomenology* (pp. 57-82). (De Gruyter).

Schaffer, Jonathan (2009). On what grounds what. In David Manley, David J. Chalmers & Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (pp. 347-383). (Oxford University Press).

Schnell, Alexander (2020). Is There a "Copernican" or an "Anti-Copernican" Revolution in Phenomenology?. In I. Apostolescu & C. Serban (Eds.), *Husserl, Kant and Transcendental Phenomenology* (pp. 391-410). (De Gruyter).

Schulting, Dennis. (2020) *Apperception and Self-Consciousness in Kant and German Idealism*. (London: Bloomsbury)

Shim, Michael K. (2017). Husserl's Spatialization of Perceptual Consciousness," *Experiential Reason: Perception, Affectivity, and Volition in Husserl's Phenomenology*. *Phaenomenologica*, eds. R. Rubio, R. Walton, S. Taguchi (Springer)

Smith, David Woodruff, & McIntyre, Ronald. (1971) "Intentionality via intensions," *Journal of Philosophy, 68*(18): 541-560

Smith, David Woodruff, & McIntyre, Ronald. (1984). *Husserl and intentionality: A study of mind, meaning, and language*. (Reidel)

Smith, David Woodruff. (2023) "Constitution Through Noema and Horizon: Husserl's Theory of Intentionality," in P. Londen, J. Yoshimi & P. Walsh (Eds.), *Horizons of Phenomenology: Essays on the State of the Field and Its Applications* (pp. 63-80). (Springer Verlag)

van Mazijk, Corijn. (2016) "Kant and Husserl on bringing perception to judgment," *Research in Hermeneutics, Phenomenology, and Practical Philosophy,* 8 (2):419-441

Zhok, Andrea. (2011) "The Ontological Status of Essences in Husserl's Thought," *New Yearbook for Phenomenology and Phenomenological Philosophy, -11*: 96-127

GOD, ABSTRACT OBJECTS, AND EXISTING NECESSARILY

Anthony Ulibarri

INTRODUCTION

There are some philosophers who support the idea that things like the number 2, the proposition that <the 4th root of 16 is 2>, and the set whose member is the set of the empty set, are abstract objects. There are also some philosophers who are inclined to hold that abstract objects like these exist. Although one can argue that there is an abundance of abstract objects, say like the property redness or the property 'being-a-car', I will restrict my argument in this paper to the kinds of abstract objects that appear to be better candidates to exist in all possible worlds like numbers, propositions, pure sets, perhaps shapes, etc. It is typically assumed that these kinds of abstract objects do not seem to depend on any concrete or contingent objects for their existence. We might say that there is a possible world where no concrete objects exist at all (no people, tables, chairs or things in general) such that you cannot gather anything together to count yet, still, there exist the true proposition <the 4th root of 16 is 2>. So, typically, the idea is that these abstract objects exist in all possible worlds or, the term I will use most often in this paper, *exist necessarily*. Peter Van Inwagen refers to these kinds of abstract objects as "free abstract objects" (Van Inwagen 2009, p. 7). I will refer to them as FAO, or sometimes just abstract objects. Covering the wide range of arguments that motivate some to believe FAO exist is beyond the scope but suffice it to say that this tends to be a prominent feature of ontology for many philosophers' metaphysical positions. For the rest of this paper, I am just going to assume that FAO exist and that these kinds of abstract objects exist necessarily.

Now, there are also some philosophers who believe in God. There are levels to theism, however. I am interested in two forms of theism (within the Christian tradition). First, is what I will call *non-creedal theism*. There is a wide range here, but the only relevant feature for the topic at hand is that this is the kind of theism that philosophically evaluates God's existence without reference to doctrines, creeds, confessions or divine revelation.[1] That is not to say that these principles are not important to the non-creedal theist, but only that their philosophical arguments do not assume them in all of their traditional interpretations, if at all. The second type is the theist

position which not only subscribes to the existence of a God, or perfect being, but also believes so within the limits (or restrictions) of creedal or doctrinal statements. Going forward, I will refer to this type of theism as *creedal theism*. This particular variety of theism will approach the nature of God and the philosophical implications that follow within the framework of such creedal statements. These creedal statements, (which as far as I can tell are held as divine revelation, self-evidently true, and axiomatic) will determine their approach to necessity and possibility as they relate to the metaphysics of a perfect being like God. Let CT refer to: *the metaphysics of a perfect being within the framework of creedal commitment*. For instance, for the non-creedal theist, it might seem reasonably clear that being all-knowing and all-powerful (or at least *more* knowing, more powerful, etc.) are the kinds of properties that one uses to describe God, or what we might call the *maximal* properties of a perfect being. On the other hand, being able to create out of nothing, being timeless, or being triune are not clearly necessary properties of God. But it gets trickier. Some philosophers who are creedal theists might assert that creedal or doctrinal commitments conflict with some proposed properties of God by other creedal theists, resulting in disputes over which properties are necessary. We might observe that the exact nature of God's knowledge about the future is unclear (at least to some) or whether God is responsible for every event that happens, and hence theories can be constructed in ways where it is equally unclear if they are consistent with CT or not. It is in this philosophical space where the topic of this paper converges concerning God and abstract objects. In other words, this paper addresses how theists working within a creedal framework, as opposed to a non-creedal framework, are attempting to resolve a philosophical issue about God and abstract objects.

Van Inwagen argues that God cannot create abstract objects (Van Inwagen 2009). He adds that FAO exist, that FAO are uncreated, and that the preceding two sentences are consistent with creedal theism. We will call this view Platonic Theism. The issue here for some creedal theists is that they claim that the proposition that can follow, namely that *there are some things that God did not create,* is inconsistent with CT. Some theists might then abandon the idea that abstract objects exist all together in favor of a more parsimonious metaphysical ontology.[2] That aside, the creedal theist Paul Gould instead argues that there is "no good reason" to think that FAO are uncreated (Gould 2014, p. 100). His motivation to argue for the conclusion stems from a statement in the Nicaean Creed (typically a statement of faith or belief within the Christian tradition) that God is the "Maker of all things, visible and invisible."[3] Gould challenges Van Inwagen's claims,

not by refuting the existence of FAO, but rather by arguing that FAO can exist, and that it is possible that God created them because for something to exist necessarily does not entail that something is uncreated.

In this paper, I will argue that Gould's solution does not work because of an ultimately incoherent notion that rejects existing eternally as a condition for existing necessarily. I conclude that existing necessarily entails existing eternally, and that there are good reasons for thinking that if something exists necessarily then it is uncreated. I then briefly address why I think these motivations stem from the particular variety of creedal theism, as I have defined here in this paper. In section 1, I will define terms. In Section 2, I cover Gould's main argument that it is logically possible that God *can* create abstract objects. In section 3, I respond, essentially challenging a central premise of Gould's argument. In section 4, I interact with the motivations behind creedal theism, and respond to a view called Modified Theistic Activism, which purports to demonstrate how God creates abstract objects. Finally, in section 5, I assess what I define as non-creedal theism, and conclude that they have less motivation, if any at all, to challenge the claim that God *cannot* create abstract objects.

SECTION 1: POSSIBLE WORLDS, EXISTING NECESSARILY AND ABSTRACT OBJECTS

First, I want to clarify how I will be using some expressions and terms. Going forward, I will be using the term 'possible worlds' frequently. When I use the term possible worlds or refer to a possible world, I simply mean 'a way that things or reality could have been'. If you happen to be reading this at home right now sipping coffee or tea, it *could* have been the case that instead you were reading this in your parked car drinking bottled water. Additionally, I might refer to something being 'impossible' or 'necessary' in a world. By this, I just mean that when evaluating possible worlds, there are 'ways things could not have been' and 'ways things must be', respectively. For example, it seems obvious that it could not be the case that there is a possible world where there are round squares—in other words it's *impossible*. Likewise, it seems *necessary* that according to the common definitions of round and square in English, that there are no possible worlds with round squares.

Another term I want to explain is what it means to 'exist necessarily'. I will adopt Gould's definition here and assume possible world semantics as "x exists necessarily = x exists in every possible world" (Gould 2014, p. 104). So, the idea here is that in every conceivable way reality

could have been, some x exists in all those realities. For instance, if we evaluate all possible worlds, a theist (creedal or non-creedal) might claim that God exists in all possible worlds, or that it's impossible that God does not exist. That would be an instance of claiming that God 'exists necessarily'. Conversely, something that 'exists contingently' is something that exists in only some possible worlds, so that 'x exists contingently = x exists in some possible worlds'.

Also, concerning abstract objects, I accept the common view that abstract objects are non-spatiotemporal, non-causal and mind independent. That is to say that FAO exist outside of space and time, they do not enter causal relations, and they would still exist in a possible world where nothing else existed (i.e. no people or minds).

Lastly, a word about the creedal/non-creedal theist distinction. The reason, I think, this is an important distinction is because the motivations to solve philosophical problems differ. I argue they differ widely. Take our topic at hand about free abstract objects. If a theist, in general, concludes in their own philosophical investigations that FAO exist necessarily and that it seems incoherent that something created them (they believe they are uncreated) then, that same theist will have motivations for either accepting that God just did not create them, or rejecting that God did not create them. If the former, then perhaps this theist is just a Platonist (someone who believes abstract objects exist) and a theist, and there are no creeds or confessions that tell them this is a problem, so they go on their happy way. If the latter, then the theist might accept Platonism, but confessions or creeds compel them to revise their Platonism because some doctrinal statement clashes with the idea of an uncreated abstract object. This is all I mean by the distinction between creedal and non-creedal theism, and I think it plays an important role here.

Now, concerning the features that, to some degree or another, remain controversial about what possible worlds and abstract objects are, or how they are structured, I leave that up to the reader. In terms of what it is to 'exist necessarily', however, this paper will explore this concept in more detail.

SECTION 2: THE ARGUMENT FOR EXISTING NECESSARILY AND BEING CREATED

Gould's argument is dependent on a broad outline presented as a case for the claim that it is possible that FAO are uncreated (Gould 2014). I have adapted this outline into the following schema:

(NC*) if God exists necessarily and God necessarily creates x, then it follows that x exists necessarily and x is created.

Based on our understanding of the term ‘exists necessarily’, NC* is meant to demonstrate that if an x exists and creates y in all possible worlds, then y both exists in all possible worlds and is created. The demonstration is then used to support the following thesis:

(NC) x exists necessarily does not entail x is uncreated.

Essentially, a conditional like the following is proposed:

(C) If NC is true, then God can create FAO.

So, the argument is that if we can establish that some object x exists necessarily and bears the property ‘created’, then it is possible that God can create FAO. Since FAO are thought of as existing necessarily and God is considered to exist necessarily, then it is possible that FAO are just created things and thus bear the property ‘created’. Finally, to further support NC*, however, Gould suggests fine-tuning our definition of existing necessarily in order to make the NC* schema more cogent. Here is the suggestion:

> “Suppose ‘x never exists’ is self-contradictory [when x exists necessarily], but ‘x sometimes does not exist’ and ‘x exists necessarily’ are consistent. Then x will exist in every possible world but will not be everlasting in every possible world” (Gould 2014, p. 106).

The idea here is that asserting that ‘x exists in every possible world’ is only asserting that ‘x is found in every possible world’, and not that there are no times in any possible world where x does not exist. So, remaining consistent with our definition of ‘exists necessarily’, the suggestion that is supposed to support NC* can be summarized as follows:

(S) x exists necessarily, and x sometimes does not exist.

Let's apply all these claims to the problem of God and abstract objects. According to Gould, S supports a consistent claim that FAO exist necessarily, but in the possible worlds where FAO exist, there might be moments or times when FAO do not exist, just before being created. If that's the case, so the argument goes, then it seems like a being like God, who exists necessarily, can choose to create FAO in every possible world. Thus, FAO also meet the conditions for existing necessarily (NC*). It appears to follow that there are no logical reasons why abstract objects are necessarily uncreated (NC), so then it's logically possible that God *can* create abstract objects (C). Thus, there is no good reason for thinking that abstract objects are uncreated.

The preceding argument is supposed to establish the logical possibility for the claim that God can create FAO. And I think the logical possibility is all Gould really needs. In other words, if we can accept S as a possible analysis of existing necessarily, then we can accept that FAO are created by God, confirming the assertion that there is no good reason to think that abstract objects are uncreated. If this is correct, then this would solve Gould's trepidation with abstract objects and the creedal commitment that motivates the dispute. That is, the apparent contradiction between the axiomatic statement that is, loosely stated, "*God created All Things*", and Van Inwagen's claim that abstract objects are not created by God. It is the "All" that is the problem.

SECTION 3: NECESSARY EXISTENCE ENTAILS ETERNAL EXISTENCE

Van Inwagen's reply to the above argument involves the non-causal nature of abstract objects. Since 'creation' is typically thought of as a causal relation, then it does not seem possible that God creates FAO (Van Inwagen 2009, p. 5). I find this reply unconvincing. If we can argue for the claim that God can create abstract objects via NC and the preceding premises, then it might not really matter that creation *is* a causal relation. If I have already established the possibility that abstract objects just are created entities, then Gould's creedal theism can contend that it just turns out that abstract objects are not causally inefficacious after all *in relation to perfect beings like God*. Otherwise, they are non-causal. Anyway, that is a possible story the creedal theist can standby, not one that I am advancing myself.

I argue rather that the relevant reply involves whether the schema NC* is a proper analysis of what it means to 'exist necessarily'. I contend that it is not, and that existing necessarily also entails existing eternally; thus principle S is a contradiction. Existing 'eternally' meets one of the

following conditions, (1) x exists eternally if x exists in all durations, moments or times, or (2) x exists eternally if x exists but there are no durations, moments or times.

First, I want to emphasize why abstract objects are non-spatiotemporal. Trivially, if an FAO exists non-spatiotemporally, then it is, by definition, non-temporal. To exist non-spatiotemporally is to exist outside of, or independent of, both space and time. So, if we accept S as a solution, we either must radically redefine what non-spatiotemporal means or just suggest that existing outside of time is existing at no times. As for the prior, I am not motivated to redefine what it means to be non-spatiotemporal. The latter makes no sense for the proponent of S because the assertion that abstract objects exist outside of time and thus exist at no times seems to suggest that, at the very least, FAO exist *eternally*, confirming (2). That is because either way, if there are no moments or times in that possible world, and yet they exist in every possible world, and it is impossible that they do *not* exist in that possible world, then they *always* exist because they exist eternally. And if abstract objects eternally exist in a possible world with no times, then they are uncreated because there are no moments before they are created, since causation is typically presumed to be a *temporal* relation and creation is a *causal* relation. So, I think even granting that FAO can enter into causal relations with God alone does not help the situation because both God and FAO are in a non-temporal state, eternally existing. Now, Gould seems to think that x existing necessarily should instead entail that x is "everlasting" (he does not mention being eternal), and everlasting is assumed to just mean "x exists throughout all periods of time" (Gould 2014, p. 105). With that definition, he argues that an FAO like the number 7 would exist at no times 't' in a world with no times, so then everlasting is not a condition of existing necessarily, hence we can still say that 'existing only sometimes' is consistent with existing necessarily. But if I am correct here, then it really does not matter whether the number 7 is everlasting or not, because if 7 is non-spatiotemporal then it is eternal, and if it is eternal and exists necessarily then, still, it *always* exists because there are no moments or durations it does not exist.

My second objection regards the direct reference of an abstract object. Suppose a creedal theist wants to insist that abstract objects are not eternal by rejecting (1) and says that existing eternally can be thought of as existing for all durations, times or moments *after* x begins to exist. Perhaps also for the sake of the argument (and I am inclined to agree with this) that a sensible analysis of existing in a possible world inevitably involves referring to times or durations, such as the claim that '7 exists at t_5'. This move just further complicates the issue because now there are

situations where abstract objects are not found in a possible world. If 7 exists in all possible worlds yet there is some subset of time t where 7 does not exist (i.e. the duration before the number 7 is created), then we would not find 7 at t_n if, say, 7 is created at t_{n+1}. So then if we reference some necessary being x in world w_1 at t_4, then the sentence 'x exists in w_1 at t_4' is simply false if x is created at t_5. The result here is also the incoherent notion that there is no number 7 from t_0-t_4. Van Inwagen mentions a similar outcome of inconceivability when asking the question of whether God can create abstract objects, specifically shapes. He comments that, "it's certainly hard to conceive of God or any being deciding whether to create a shape. If the shape doesn't exist, what is it he's deciding whether to create?" (Van Inwagen 2015, p. 289). As far as I can tell, the same goes for the number 7. What is it that God is deciding to create when they create the number 7 at t_5?

There is also another reason for rejecting S. There is the claim that it is just flat-out part of the definition of 'existing necessarily' when considering that possible worlds are also possible situations. Timothy Williamson notes the following:

> Something exists necessarily only if it exists in all possible situations; but all past, present and future situations are possible, so it exists in all past, present and future situations, so it exists eternally. (Williamson 2002, p. 237).

So far, we have assumed that existing necessarily is existing in all possible worlds, which I broadly defined as 'a way that things or reality could have been'. I will refer to this broad sense of possible worlds from now on as the *broad-pw-interpretation*. But other interpretations also define a 'possible world' in a narrow sense as including *possible situations* (Stalnaker 1986). Indeed, all possible situations include past, present and future situations, and anything existing necessarily would exist in all those situations. In other words, they would exist *eternally*. I will call this the *narrow-pw-interpretation*. Should we accept this analysis? My intuition tells me that the narrow-pw-interpretation was already recognized in the previous objections. The kinds of things that are assigned necessary existence are not the kinds of things evaluated simply in the broad sense. They are the kinds of objects that are evaluated in the narrow sense as well. Consider the following claim: "Lebron James exists necessarily, and it is false that 'LBJ exists' before he is born and true after he is born". This claim is false, but it is coherent. As a matter of fact, assuming we all agree that LBJ is not eternal, the following conditional is true: "If it were the case that LBJ existed

necessarily, then the sentence 'LBJ exists' would be false before he was born and true after he is born". But is this conditional true because it is false that 'LBJ exists necessarily', so anything follows, or is it because LBJ actually does exist necessarily? We know that LBJ (unfortunately) does not exist necessarily in either the narrow or the broad sense of possible worlds. But if he did, and we reject the narrow-pw-interpretation, then it seems like the consequent is true. That is because rejecting the narrow-pw-interpretation, we are not evaluating Lebron James in terms of possible situations. We are not asking questions like, "Would LBJ exist in all past, present and future situations?" Accepting only the broad-pw-interpretation we lower the bar when assessing whether some being, or object exists necessarily. One of the reasons that the number 7 turns out to exist necessarily in our analysis is because 7 is prime in all past, present, and future situations; and unlike LBJ, these properties are true of the number 7 in all possible situations. More importantly, I think the best explanation for everything we have considered is that existing necessarily entails existing eternally.

SECTION 4: A CREEDAL THEISTIC MOTIVATION

So, do the motivations for creedal theism cause one to outright deny existing eternally as a condition for existing necessarily? Well, the answer varies. As I mentioned at the beginning of this paper, our topic converges as a dispute between two professed creedal theists who maintain that both of their views are consistent with creedal axioms. Considering Gould's view, the motivations are quite clear:

> Theistic philosophers who wish to endorse Platonism face a dilemma. Either affirm uncreated abstract objects and risk theological unorthodoxy with respect to the doctrines of creation and divine aseity or somehow bring the Platonic horde safely within the bounds of theological orthodoxy but risk incoherency. (Gould 2014b, p. 283)

The "dilemma" referenced here is the issue that abstract objects are typically understood as being uncreated, and the risk of "theological orthodoxy" mentioned is the apparent inconsistency with the axiom that God created "*All* things". However, Van Inwagen, seems willing to suggest an alternate reading of the axiom. He comments:

> But must a Christian take this creedal statement to mean that God has created everything, everything tout court, everything simpliciter, everything full stop, everything period? Or is it permissible for the Christian to regard the range of the quantifier 'everything' in the sentence 'God has created everything' as restricted to a certain class of objects… (Van Inwagen 2009, p. 3)

So, Van Inwagen avoids the dilemma altogether by proposing an alternative reading to the quantifier "All" in the axiom. From my point of view, this strategy seems just fine to me but then again, I do not have much stake in the debate as to whether a doctrinal confession is properly interpreted or not.

What I will say is that Gould maintains the strong conviction that God absolutely creates "All" things and that this includes the Platonist horde of abstract objects. His view called *Modified Theistic Activism* (MTA) sets out to rationalize what it would be like for God to create abstract objects like FAO. The view claims that abstract objects basically inhabit and "exist in two realms: the divine mind and Plato's heaven" (Gould 2014c, p. 61). The basic idea is that some abstract objects, particularly FAO in the way we have discussed here, are constituents of God's mind. The more complex idea is that God's mental activity somehow causes FAO because, according to MTA, God is *logically prior* to something like the number 7, or the proposition that <7 is prime>. Logical priority in this case is meant to be understood as being a non-temporal causal explanation (Gould 2011, p. 137). That is, we might say, disregarding the temporal order of events, that, in part, the ball is the explanation of why the window is broken. So, is this view consistent with necessarily and eternally existent objects? I do not think so.

My contention is that MTA does not gain much traction unless Gould's initial premise that 'x exists necessarily, and x exists sometimes' is consistent, and it is not. Interestingly, Gould has it that our candidates for FAO (number, propositions, pure sets, etc.) result from the "creative activity" of God's mind but that properties (not of God) and relations are the inhabitants of Platonic heaven, existing "wholly apart from God" (Gould 2014b, p. 284). I think the motive for this dichotomy is that it seems incoherent that some abstract objects, like the number 7, are created simpliciter, whereas possibly the property of 'being an analytic philosopher' is created in the sense that the property absolutely did not exist and then it did at creation and post-creation. So, my guess is that the abstract objects in Platonic heaven that are wholly apart from God are thought of as

easier to explain for MTA. But what about the property 'thinking about the number 7"? If this property is not part of God's mental activity and is absolutely created at some later instance, then does that mean God never thought about the number 7 prior to the property's creation? If that's true, then there's no explanation for God's mental activity about the number 7. What I find intriguing is that FAO are not even the chief factor for rationalizing that some abstract objects are in God's mind, and some are not. The primary motivator is God's own divine properties, such as 'being divine' (Gould 2011, p. 132). The thought is that if the property exemplification of God (like God exemplifying being divine) can be explained away, then the mental constituents of God's mind, like FAO, get explained with it. But I find it the other way around. If we cannot explain how God lacks the property 'thinking about the number 7' when the abstract object 7 is supposed to be in God's mind already, then we certainly cannot explain how God exemplifies his own divine properties (if we want to insist that God creates all of these abstract objects).

Regarding logical priority, the argument goes that God would be logically prior to the number 7. Perhaps I can just submit that God is somehow logically prior to his being divine, but that certainly does not explain how the number 7 is logically subsequent to God. If the number 7 exists eternally because it exists necessarily, then claiming that the number 7 can be explained by God via logical priority is just false. Since, the number 7 is prime in all past, present, and future situations, no such logical priority is required to explain that '7 is prime' is true. A counterfactual form can be offered here to better explain: If God does not exist necessarily, then '7 is prime' is true. I am certainly not asking the reader to accept that the antecedent is true. But if it were true, and Platonism is the presumed view, how could the consequent be false? It seems to me like the number 7 requires no explanation since it exists necessarily. That existing eternally is built-in to existing necessarily, which is the outcome of the kinds of objects we determine as existing necessarily.

SECTION 5: A NON-CREEDAL THEISTIC MOTIVATION

In my view, a condition for being a non-creedal theist is not that she considers creeds or confessions false. The only condition is that the creeds or confessions play a lesser role or none at all in their philosophical inferences. The 'lesser role' I'm describing is quite similar to Van Inwagen's reinterpretative suggestion for an alternate reading of creedal axioms.[4] If creedal

confessions play no role, then it goes without saying that they do not motivate where their philosophy takes them.

Consider Nagasawa's description of, what we will call, the *Anselmian thesis* (AT) that "God is that than which no greater can be thought" (Nagasawa 2013, p. 3). Although many theists, creedal or non-creedal alike, might endorse AT, the intention is that this approach is purely metaphysical insofar as it reflects on the properties that would be a being great. For example, a basic linear model of AT can be thought of in the following way: if you laid all beings and objects, possible and actual, in a straight line, from least to greatest, then God occupies the very first spot in line. That is to say that they are the top link in the great chain of being. The argument then establishes the great-making and maximal properties of God. So, a property is 'great-making' if it's better that God has it than not having it when bearing in mind that they are the greatest being. A maximal property would then be those properties that God bears once we have established that they are the greatest. In the traditional sense, those would be the omni-properties such as omnipotence, omniscience, omnibenevolence, etc.

However, as Nagasawa notes, the omni-properties of God are not resistant to scrutiny, and some philosophers might question whether it is even possible "for any being to be omniperfect" (Nagasawa 2008, p. 588). There are plenty of reasons for this that I will not get into, however it stands to reason that if abstract objects are the types of entities that exist necessarily, then it is just unfeasible to think that God can create them anymore than God can create a round square. It follows that there is no reason to attribute God maximal power in the form of an unqualified omni-property that relies on the idea that all entities, both necessary and contingent, are ontologically dependent. In Nagasawa's model, the Anselmian thesis does not entail an 'omniGod' (ibid., p. 591). That is, a God whose properties must be conceived of in the limitless sense.

So, is it consistent with a confessional framework (CT) to admit that God is not omnipotent? That depends on whom you ask. The focus here is the non-creedal theist, so the critical decision is whether denying omnipotence should follow from denying that God can create abstract objects. In a sense, the only thing we refrain from doing is forcing the maximal interpretation of what it is to be 'All-powerful'. In my view, even if God is not omnipotent in the traditional sense, then they are at least more powerful than everything else in the universe. God would even be more powerful than the abstract objects they are not able to create, since we assume that FAO are causally inert. The metaphysical fork-in-the-road is whether one considers any of

these ideas as contradictory to the 'plain meaning' of the creeds or confessions. A non-creedal theist might not care, or at least have no issues with reading creeds that were written centuries ago with a modern lens. I think that this makes the original *meanings* irrelevant yet maintains *cultural* relevance. I digress, but it is most certainly the case that the creedal theist would be concerned with both the plain meaning *and* the original meaning. That being said, Van Inwagen's suggestion is suspect to me for the very reason that if we were to ask the original authors of such confessional documents what they meant when they wrote these axioms that they would not be amused in the least with the idea that are some things that God did *not* create.

My point here is that there are alternative conceptions of God that are purely philosophical in nature.[5] Some of these conceptions have an impact on how creeds or confessions are interpreted, or they have no impact at all since creeds or confessions were never referenced when establishing God's nature in the first place. I contend that both claims comprise the activity of the non-creedal theist, and that they lack any real motivations for challenging the notion that God cannot create abstract objects. So, in my view, the creedal theists alone, as I have defined them, are susceptible to the arguments in this paper and Platonism is inconsistent with only this variety of theism.

CONCLUSION

If there is a major takeaway from what has been presented here, it is that the concept of existing necessarily entails existing eternally. I aimed first to present a theist's argument that it is logically possible that God can create abstract objects. I then challenged a major premise of this argument, effectively contending that we cannot make sense of non-eternal necessarily existing objects. I concluded that Gould's two claims in his major premise that 'x exists necessarily, and x exists sometimes' are inconsistent without a reinterpretation of what existing necessarily is. Yet this reinterpretation is not effective and does not explain but rather complicates our understanding of free abstract objects. Finally, I responded briefly to a view called *Modified Theistic Activism* maintaining that the view does not address the issues involved without first addressing whether it is even logically possible that God can create FAO. Since the view does not address those issues, it is not a valid option. I then covered a non-creedal theistic model, where I determined that the claim that 'God cannot create abstract objects' is primarily motivated by a creedal theist's conviction to hold to certain axioms that are part of their tradition. Unless they are willing to either

reinterpret these creeds or abandon them as philosophical premises Platonism, as conventionally understood, will be inconsistent with their position.

Notes

1. William Craig's holds a *nominalist* ontology, found in: William Lane Craig, "Anti-Platonism," in *Beyond the Control of God? Six Views on the Problem of God and Abstract Objects*, ed. Paul M. Gould (New York: Bloomsbury, 2014), p. 115.

2. The spectrum here for 'non-creedal' involves theists that consider themselves as part of traditional Christianity and those who consider themselves as part of alternative or progressive forms of the Christian tradition.

3. In mainstream Christianity, the Nicaean Creed is considered to be a fundamental statement of belief and is a profession of faith for both Orthodox and Catholic traditions.

4. As already noted, Van Inwagen would consider his view as what I have defined as a 'creedal' version of theism. My remark is simply that a non-creedal theist might engage in the same activity without an ultimate concern as to whether their conclusions align with mainstream traditions or not.

5. Philosophers might argue that what I define as 'creedal theism' is very much a philosophical practice as well. While I agree, my distinction at this point is meant to convey that a creedal axiom is treated much like axioms in general. That is, the kind of claim that is self-evidently true. However, this is not all obvious for either the non-Christian or possibly philosophers on one side of the non-creedal spectrum since what is self-evidently true to a creedal theist derives from holy texts that are themselves considered self-evidently true and plainly interpreted (at least in cases where the axioms are derived). So, the idea of what is self-evidently true can vary widely between a non-creedal theist and a creedal theist.

Bibliography

Botefuhr, W. D. C. (1885) *Nicene Creed* (Fort Smith: Boetefuehr, W. D. C, 1885. Notated Music) Retrieved from the Library of Congress, <www.loc.gov/item/2023859387/>

Gould, Paul (2011) "Theistic Activism, a New Problem and Solution," *Philosophia Christi* 13, pp. 127-139

——(2014) "Can God create Abstract Objects? A reply to Peter Van Inwagen," *Sophia* 53, pp. 99-112

——(2014b) "Theistic Activism and the Doctrine of Creation," *Philosophia Christi* 16, pp. 283-296

——Davis, Richard Brian (2014c) "Modified Theistic Activism," in: *Beyond the Control of God? Six Views on the Problem of Abstract Objects*, pp. 51-79 (New York: Bloomsbury Press)

Nagasawa, Yujin (2008) "A New Defence of Anselmian Theism," *The Philosophical Quarterly* 58, pp. 577-596

——(2011) "Models of Anselmian Theism," *Faith and Philosophy: The Journal of the Society of Christian Philosophers* 30, pp. 1-25

Stalnaker, Robert. (1986) "Possible Worlds and Situations," *Journal of Philosophical Logic* 15, pp. 109 - 123.

Van Inwagen, Peter. (2009) "God and Other Uncreated things," in: Timpe, Kevin (ed.), *Metaphysics and God, Essays in honor of Eleonore Stump*, pp. 1-19 (Routledge: Taylor & Francis)

——(2015) "Did God create shapes?," *Philosophia Christi* 17, pp. 285-290

Williamson, Timothy (2002) "Necessary Existents," in: O'Hear, Anthony (ed.), *Logic, thought, and language*, pp. 233-251 (Cambridge University Press)

A DEFENSE OF THE "NO FALSE GROUNDS" ACCOUNT FOR KNOWLEDGE

AbdulRahman Alzarouni

INTRODUCTION

Some philosophers have traditionally held that knowledge is defined as 'justified true belief.' The definition can be stated as follows: S (where S is any subject) knows that P (where P is a proposition) if and only if: a) S believes that P, b) P is true, and c) S is justified in believing that P. This definition has been criticized by Edmund Gettier (1963) (I will call those criticisms: "the Gettier case/cases") in his famous paper, "Is Justified True Belief Knowledge?" One of the responses that has been proposed to save the traditional definition from the criticisms suggests that a fourth condition must be added along the three traditional conditions. There have been different accounts of adding a fourth condition for knowledge, but in this paper, I am defending the "No False Grounds" account for knowledge, which adds the following condition: d) S's belief that P is fully grounded. I am defending it against several counterexamples that try to show that the "No False Grounds" account fails to provide a good account for knowledge. I need to clarify some points before I get into the discussion.

First, I need to show how definitions work in general. There is a left side and right side to a definition. What separates both sides is the phrase 'if and only if'. Both sides must be true or false for the definition to be true. We don't want to have a case where one side is true while the other is false. So, for example, we can't have a case where a person knows a proposition P without satisfying all the conditions on knowledge (assuming that the conditions are a) S believes that P, b) P is true, and c) S is justified in believing that P). Also, we can't have a case where a person satisfies all the conditions for knowledge but does not know proposition P. Gettier's criticisms show that a person satisfies all the conditions but fails to know that P. Secondly, I want to emphasize that the kind of knowledge I am concerned about in this paper is propositional knowledge. Propositional knowledge has the following form: S knows that P, where S knows that such and such. However, I am not concerned with knowledge by acquaintance, which has the form:

S knows x (where x is a specific person). I am concerned about what is attributed or said about a person x. I am concerned about asserting a certain fact about a person, place, animal, event … etc.

This paper will be divided into three sections. In section one, I will defend each condition of the traditional definition of knowledge and explain why each is a necessary condition. I will also introduce a Gettier case and show how it challenges the traditional definition of knowledge. In section two, I will show the structure of the Gettier case and argue for the "No False Grounds" account as a solution to the problem. In this section, I will also introduce other Gettier-style cases that were offered by other philosophers. Those cases challenge the "No False Grounds" account by showing that one side of the definition could be true while the other is false. Finally, in this section, I will try to show that those cases do not refute the "No False Grounds" account. In section three, I will formulate a modified version of the "No False Grounds" account by modifying the justification condition and modifying the fourth condition from how it was originally stated while the first two conditions remain the same so it could avoid the counterexamples. My modification is as follows: c) S justifiably believes that P, d) S's belief that P is not based on any falsehood, whether it is implicit or explicit.

SECTION 1

When I say that 'I know that it is raining right now in Chicago', is that the same thing as saying that 'I believe that it is raining right now in Chicago'? To believe something does not guarantee knowledge. Two people could hold opposing beliefs. I could believe that it is raining right now in Chicago and my friend could believe that it is not raining right now in Chicago. The content of each belief (whatever comes after the 'that' clause) must be either true or false. If the content is true, then its negation can't be true at the same time. So, from a logical point of view, we can't say that belief is knowledge because we can't have two contradicting views count as knowledge, but belief is a necessary condition for knowledge. What I mean by this is that if we know that P, then we believe that P but not vice-versa. We need belief as a necessary condition because it is some kind of conviction. We don't want, for instance, to imagine that P or thinking that P is a necessary condition for knowledge because we could imagine or think without holding any conviction toward P.

When we believe that P, we usually assume that P is true; otherwise, it would make no sense why we believe that P. We don't say that we believe that P but assume that it is false. I could

for instance say, ‘I believe that it is not raining right now in Chicago’. Even if the content of my belief has a negation, that doesn’t mean that I assume that my belief is false. Thus, I believe that the negation of P would make not P true and P false. Now, I will give a simple theory of truth that serves what I am going to argue for in this paper. When I say, ‘it is raining right now in Chicago’, I use that proposition to correspond to the state of affairs of the world. So, if it is actually raining in Chicago right now, then the proposition ‘it is raining in Chicago right now’ would be true. So, in this case, the proposition would correspond to the state of affairs of the world. On the other hand, if it is actually not raining in Chicago right now, then the proposition ‘it is raining in Chicago right now’ would be false. So, the proposition would not correspond to the state of affairs of the world. Now, let us consider why truth is a necessary condition for knowledge. It is a necessary condition because it would be absurd to say that I know that P, but it is false that P. For example, it would be absurd for me to say that ‘I know that there is a table in front of me but it is false that there is a table in front of me’. To know some proposition P requires truth in all cases.

At this point, we have established that belief and truth are necessary conditions for knowledge. But are they jointly sufficient for knowledge? According to the traditional definition and the account I am defending, the answer is no. Consider this case, suppose my friend John tells me that my brother is in the shopping mall. Let’s grant that John believes that my brother is in the shopping mall and my brother is actually in the shopping mall. When I ask John how does he know this? He answers, “I just believe it.” In this case, John is lucky, he just believes that my brother is in the shopping mall without having reasons for it. We don’t want knowledge to involve cases where someone believes that P due to luck, where P happens to be true rather than reasons. So, we need a third necessary condition for knowledge and that condition is justification.

Now, I will show what counts as ‘justification’ or count as ‘strong justification’. We could believe that P for the wrong kind of reasons, and I want to rule out those cases as being cases of justification. Suppose that I am sick, and the doctor prescribes a medicine for me and tells me that if I take the medicine, I will probably regain my health in about a week. Let’s assume that the doctor has diagnosed me correctly and that the medicine is effective in my case. Let’s assume that those are strong reasons for me to believe that I will regain my health in about a week. Suppose that I truly believe that I will regain my health in about a week because I always like to believe whatever my wife tells me to believe because it makes her feel happy. In this case, I believe truly that P for the wrong kind of reasons. Those are not the kind of reasons I am calling ‘justification’.

But there is something interesting in this case. I had strong reasons at my disposal, but I did not consider it. My belief was not based on those strong reasons. If I had used those strong reasons, I would have had knowledge according to the traditional account. So, does that mean I was not justified in this case? I would say yes in this case. But later in the paper, I will present a case where a person has justification at his disposal but doesn't use it, yet the person is justified.

So now let me say a little bit about what could count as strong reasons. Strong reasons are reliable reasons. I am using 'reliable' in a weak sense. Let's say that reliable reasons are like perception, testimony of people who don't lie often, and predictions that have high probability for preserving truth, etc. Let us say that in the cases that I am considering, there is consensus that those reasons are strong reasons. I will provide a case to show what could count as strong reasons versus weak reasons. Suppose Sara wakes up in the morning and she goes outside and meets a bunch of people claiming that it is raining in Chicago right now because the clouds are grey over there. Suppose Sara believes that it is raining in Chicago right now for the reason that the clouds are grey over there. Suppose that it is actually raining in Chicago right now. In this case, Sara doesn't know that it is raining in Chicago right now because her reasons are weak. Her justification condition is not satisfied in this case. The greyness of the clouds is not a good indicator for the fact that it is raining, or it is going to rain. Now consider the same case again. But this time, Sara watches two weather channels that says that it is raining in Chicago right now and sees that it is raining in Chicago right now via a live camera. In this case, the justification condition is satisfied for Sara. There are different theories of justification that argue for what counts as strong justification. For the purposes of this paper, I think the examples of justification I provide would suffice, and if anyone reading this paper is not satisfied with them, alternative scenarios could be imagined.

The three conditions I have outlined are necessary, but not jointly sufficient conditions for knowledge because Gettier cases pose problems. Suppose that Smith and Jones both work at the same office. Smith knows that Jones is a hard worker and has heard from the boss that Jones would get a raise. It also happens that Smith has seen Jones count his ten coins and put them in his pocket. Given all this information, Smith forms the following belief: Jones will get a raise and he has ten coins in his pocket. Smith deduces the following from the previous belief: The man who will get a raise has ten coins in his pocket. But unknown to Smith, Smith (himself) has ten coins in his pocket and Smith (himself) will get a raise. Also, the boss misspoke when he said that Jones will get a raise. This case shows that Smith truly believes that the man who will get a raise has ten coins

in his pocket. Smith is also justified in believing so because he has strong reasons in the first place to believe "Jones will get a raise and he has ten coins in his pocket" and Smith deduces "the man who will get a raise has ten coins in his pocket" on the basis of "Jones will get a raise and he has ten coins in his pocket". So, this turns out to be a case where the three conditions are satisfied while Smith lacks knowledge. Smith lacks knowledge because he thinks Jones will get a raise and Jones has ten coins in his pocket. It seems like Smith got lucky that the statement "the man who will get a raise has ten coins in his pocket" truly picks him up.

SECTION 2

Let us now consider the structure of the Gettier case. First, Smith has gathered evidence, and this evidence is the justification of his belief that Jones will get a raise and he has ten coins in his pocket. Secondly, Smith's belief that Jones will get a raise and he has ten coins in his pocket is false. Finally, Smith truly deduces 'the man who will get a raise has ten coins in his pocket' from the false statement 'Jones will get a raise and he has ten coins in his pocket'. The Gettier case shows that Smith truly believes 'the man who will get a raise has ten coins in his pocket' and is justified in believing it. Now we could ask the following: How is Smith justified in believing 'the man who will get a raise has ten coins in his pocket'? If we consider the structure of the Gettier case, we notice that there are two steps of justification. The first step of justification is when Smith gathers evidence about Jones, e.g., Jones is a hard worker, the boss's claim that Jones will get a raise, and when Smith counts the coins that Jones put in his pocket. The evidence Smith gathers justifies his belief that 'Jones will get a raise and he has ten coins in his pocket'. The second step of justification is when Smith uses the statement 'Jones will get a raise and he has ten coins in his pocket' to deduce 'the man who will get a raise has ten coins in his pocket'.

Someone could raise an objection against the Gettier case by saying that Smith lacks knowledge because the justification condition is violated. So, the Gettier case is not a case where all the three conditions are satisfied while subject S lacks knowledge. Someone could say that Smith is not justified in believing that 'the man who will get a raise has ten coins in his pocket' because in the first place, we have assumed that Smith has a justified false belief. We assumed that Smith is justified in believing the false statement 'Jones will get a raise and he has ten coins in his pocket'. Someone could object by saying that the 'justified false belief' is causing trouble for Smith to know that 'the man who will get a raise has ten coins in his pocket'. So, we can't have

justified false beliefs for knowledge. This objection also shows that Smith can't deduce 'the man who will get a raise has ten coins in his pocket' from 'Jones will get a raise and has ten coins in his pocket' because it is deduced from a justified falsehood. I don't think that those are good objections against the Gettier case. It seems that there is nothing wrong with having a justified falsehood. In our daily life, it seems that we could really have strong evidence to believe a certain proposition, but it turns out that our belief is false. As I have indicated earlier in this paper, we are assuming that 'justification' counts as strong reasons. Suppose that it was true that Jones will get a raise and he has ten coins in his pocket. The justification seems to be the same as in the case where Smith has a false belief about Jones. So, if we want to say that Smith is not justified in believing that 'Jones will get a raise and he has ten coins in his pocket', we must say that Smith is also not justified in the case where the belief about Jones is true. The same applies to the true belief that 'the man who will get a raise has ten coins in his pocket'. If Smith is not justified in holding the belief 'the man who will get a raise has ten coins in his pocket' for the reason that it is deduced from a justified falsehood, then Smith is also not justified in holding the same belief in the case where the belief 'the man who will get a raise has ten coins in his pocket' is deduced from a justified truth. The evidence that Smith relies on in forming the belief 'Jones will get a raise and he has ten coins in his pocket' is the same in the case where the belief is false and in the case where the belief is true.

So, the Gettier case still shows that Smith lacks knowledge even though the three conditions are satisfied. Now I will introduce an attempt that tries to respond to the Gettier case by adding a fourth condition for knowledge. An early version of this account was introduced by Michael Clark (1963). Clark says the following: S knows P if and only if a) P is true, b) S believes that P, c) S is justified in believing that P, and d) S's belief that P is fully grounded. So, what is the significance of adding this fourth condition? First, let me explain what it means for a belief to be fully grounded. We have seen so far that the structure of a Gettier case shows a subject S having a justified false belief among his reasons. In some cases, one of S's evidence for believing P comes from listening to someone claiming to give a fact. But according to Clark, the person giving the fact must not just say it randomly. The person giving the fact must be reliable. Secondly, and most importantly, there should be no false proposition involved in the steps that S takes in forming a justified true belief. We have seen that there is a false proposition that Smith believes, which is 'Jones will get a raise and he has ten coins in his pocket'. So, the Gettier case would violate the

fourth condition for knowledge. Clark's account requires that there must be no false ground/proposition in the reasoning process that leads to the justified true belief conclusion. But according to other Gettier-style cases, Clark's account doesn't seem to work.

Now, I will present four counterexamples that challenge the "No False Ground" account for knowledge. The first counterexample is similar to Richard Feldman's *Alternate Route* case. Suppose I work with Jones and Smith at the same department. I saw Jones count his ten coins and put them in his pocket. Suppose that someone in the department has mistakenly sent a document to Jones's office. This document says that Jones will get a raise. As soon as Jones receives the document he rushes to my office and shows it to me. Also, Smith receives a document that says that Smith will get a raise. It also happens that Smith has ten coins in his pocket. In fact, Smith is the one who will get a raise and not Jones. Suppose I form the following belief: Someone in my department will get a raise and he has ten coins in his pocket. Let us assume that there is only one person who will get a raise and all who work in my department are males. The *Alternate Route* case shows that I could reason as follows: 1) Jones who is in my department has a document that says that he will get a raise, and Jones has ten coins in his pocket; 2) someone who is in my department has a document that says that he will get a raise, and he has ten coins in his pocket; and 3) someone who is in my department will get a raise and he has ten coins in his pocket. Let me show the difference between the *Alternate Route* and the typical Gettier case. In the typical Gettier case, my reasoning would be as follows: 1a) Jones who is in my department has a document that says that he will get a raise, and Jones has ten coins in his pocket; 2a) Jones who is in my department will get a raise and he has ten coins in his pocket; and 3a) Someone who is in my department will get a raise and he has ten coins in his pocket. The only difference between the two cases are the steps I make at 2) and 2a). The first step in the two cases, in 1) and 1a), is my evidence, hence, my justification for 2) and 2a). The second step in the two cases shows that 2) is true while 2a) is false. So, 2) is a justified truth while 2b) is a justified falsehood. The third step in the two cases shows that both 3) and 3b), which are true, are inferred from 2) and 2a) respectively.

The steps I make when I reason in the *Alternate Route* are all true leading to the true conclusion that 'someone who is in my department will get a raise and he has ten coins in his pocket'. All the conditions for knowledge seem to be satisfied in this case. I believe that someone who is in my department will get a raise and he has ten coins in his pocket. It is true that someone who is in my department will get a raise and he has ten coins in his pocket. I am justified in

believing that someone who is in my department will get a raise and he has ten coins in his pocket. There is no false proposition that my belief that 'someone who is in my department will get a raise and he has ten coins in his pocket' is based upon. Or we could say that there is no falsehood among my reasons for believing the true proposition. So, there is no false ground that my belief is based upon (the steps I take to reach my conclusion are all true). But I do not know that someone in my department will get a raise and has ten coins in his pocket. I do not know it because it seems that my belief is somehow lucky. I would say that my belief in the conclusion would be lucky if there is a false proposition somewhere along the steps I take to reach my conclusion. In the *Alternate Route* case, there is a false step that I make in reaching my conclusion, but that step is implicit. I am not sure whether Clark's fourth condition was meant to show that even implicit propositions/premises are counted as grounds for reaching a conclusion. My interpretation of the "No False Grounds" account would include those implicit steps as well. So, in the *Alternate Route* case, I do not know that someone who is in my department will get a raise and he has ten coins in his pocket because there is an implicit false step in my reasoning. The implicit false step in my reasoning violates that fourth condition which says that S's belief that P must be fully grounded. But what is the implicit false step, someone might ask? The implicit false step is the belief that Jones who is in my department will get a raise and he has ten coins in his pocket. If someone were to ask me, "Do you believe that someone in your department will get a raise and has ten coins in his pocket because your next-door neighbor will get a raise and has ten coins in his pocket?" I would probably say "no." My answer would probably be "I believe it because Jones will get a raise and he has ten coins in his pocket." I will say this because I am not even justified in believing that my next-door neighbor will get a raise and he has ten coins in his pocket. So, the *Alternate Route* case does not refute the "No False Grounds" account for knowledge.

Let's consider another counterexample. The second counterexample is similar to an example given by Roderick Chisholm (1977). I will call it the *Robot-Sheep* case. Suppose they are about to film a movie and they have chosen to shoot a scene of a shepherd walking with a sheep in a vast field. For ethical reasons, they decided to bring a robot that looks and behaves like a sheep and is indistinguishable from a real sheep. It also happens that the filming location is near my house, and I happen to pass by and see the sheep. I also look at it carefully as I would have looked at a real sheep. Suppose that after I pass by the field, I meet my friend and he says, "This field is so green! It would be nice if there were sheep in this field." I reply to my friend and say, "There is

a sheep in the field." Suppose that there is actually a real sheep walking in the field which I did not see. This case is supposed to be a bit different from the typical Gettier cases because it involves perception as my justification for believing that there is a sheep in the field. It is argued that I directly form the belief that there is a sheep in the field given my perception. Since I truly believe that there is a sheep in the field, I am justified in doing so because of my perception, and I do not infer the conclusion 'there is a sheep in the field' from a false proposition, then I satisfy all of the four conditions. Yet, I do not know that there is a sheep in the field because it seems like I was lucky in truly believing that there is a sheep in the field. In this case, I will argue that I was lucky because there is an implicit false step in my reasoning.

Let's consider how my reasoning went. First, I see something that looks like a sheep and I believe that it is a sheep. My perception of that thing is my justification for my belief that there is a sheep in front of me. Finally, I truly believe that there is a sheep in the field. I truly believe that there is a sheep in the field on the basis of my belief that I see a sheep in front of me. What I see in front of me is not sheep but a robot. Let me make this clearer. Suppose that the robot-sheep has a pink bell tied around its neck. Let's suppose that the real sheep which is in the field as well has a blue bell tied around its neck. Suppose that after an hour I return to the field with my friend, and I see the real sheep walking by. My friend tells me, "Look, there really is a sheep in the field! You were right in saying that there is one." When I look at the sheep I say, "Well, I did not mean that sheep; there was another sheep that had a pink bell tied around its neck." So, when I said that there is a sheep in the field, I actually meant the sheep that has a pink bell tied around its neck. My assumption that the thing that I see that has a pink bell around its neck is a sheep is false. Although I did not explicitly infer the conclusion 'there is a sheep in the field' from 'I see a sheep in front of me', still, my perception of the robot-sheep is about a particular object. It could be put as follows in terms of a statement: S) that animal in the field is a sheep. S) is false. So, the *Robot-Sheep* violates the fourth condition for knowledge. There is a false ground in my reasoning for the conclusion 'there is a sheep in the field.'

Let's consider another counterexample. The third counterexample is similar to an example given by Alvin Goldman (1976). I will call it the *Fake Barn* case. Suppose Sara is driving along the road with her friends and passes by a landscape. Just for fun, she looks at different objects and tries to identify those objects by calling its names and she does so correctly. Suppose that she sees a barn and says, "That's a barn." Now the question is, does she know that what she sees is a barn?

Let's assume that she has good vision, the sky is clear, she knows how a barn looks like from her experiences, and it is actually a barn. I think most of us would say that she knows that the thing that she sees is a barn. But now consider this. Suppose that the landscape is full of barn facades (fake barns) that look exactly like real barns if someone were to look at them from the front side. But those barn facades have no back walls, side walls, and no interior. Let's also suppose that Sara has not yet encountered those fake barns. Would we still say that Sara knows that what she sees is a barn? Goldman argues that Sara does not know that what she is seeing is a barn. She does not know because she is lucky. If she encounters a fake barn or if the barn she sees was in fact fake, she would probably believe that it's a barn. This case also challenges the "No False Grounds" account in showing that Sara's perception immediately justifies her belief that she is seeing a barn. In this case, Sara truly believes that what she sees is a barn, perception is her justification for that belief, and her belief is not inferred from any false proposition. So, all four conditions are satisfied, yet Sara does not know that what she sees is a barn.

In the *Fake Barn* case, at least as Goldman presents it, I can't say whether Sara knows or not because the picture is not clear enough. How is it not clear? Let's consider the true conclusion 'Someone who is in my department will get a raise and he has ten coins in his pocket.' If someone asks me, "How do you know that?" I could immediately say, "Because Jones who is in my department will get a raise and he has ten coins in his pocket." Now let's consider one possible answer that Sara would give if she were asked, "How do you know or why do you believe that what you are seeing is a barn?" Sara could say that she knows that it's a barn because it looks like a barn. But when we say that something looks like something else, we need a sample of at least another object that resembles the object in question, which is the barn that Sara is currently seeing or is referring to. To say I know something just because I see it seems a bit unclear. Those are possible answers that Sara could give, and each answer would bear on her knowledge claim. Answer 1: The barn I am seeing looks like the barns that are located near my house (those are real barns). Answer 2: The barn I am seeing looks like the barns over there (the barn facades). So in 1), the reasoning could go as follows: a) What I see has such and such features which resemble the features of the barns that are located near my house; b) therefore, what I see is a barn. In 2), the reasoning could go as follows: a1) What I see has such and such features which resemble the features of the barns over there (the barn facades); b2) therefore, what I see is a barn.

In both cases, Sara believes that what she sees is a barn and her belief is true, but in a1), the proposition is false because there is a presupposition that the objects over there (the barn facades) are real barns. So, b2) is inferred from a false proposition. In all cases of direct perception, if someone wants to justify their claim about identifying a certain object that it is such, then there must be a sample or model of an object that is used as a justification as the starting point in the chain of reasoning and this model must be self-justifying. The remembrance of the model would also suffice as justification. Or, the model could be a concept. (Let's stick with objects rather than concepts in this case; I don't want to discuss concepts for various reasons.) Let's consider the barns that are near Sara's house as the model she uses to justify the fake barns she sees (the fake barns look like the model). So, when Sara now refers to the fake barns and says, "What I see are barns," she has a justified false belief. So in b2), Sara's true belief that what she sees is a barn is inferred from a justified falsehood. I would say that in this case Sara does not know because the fourth condition is violated (there is a false proposition amongst the steps of reasoning that leads to the conclusion). Suppose that every time Sara sees a real barn and believes that what she sees is a real barn on the basis of comparing it to the model, then Sara would know that what she sees is a real barn. So, in the *Fake Barn* case, Sara would know that what she sees is a barn if she uses the model as a basis of her justification. Now consider if there were no fake barns on the landscape but all real barns, and I happen to see a barn and belief that what I see is a barn. Let's also assume that I have not yet encountered the other real barns on the landscape. Would I know that what I see is a barn? I would say yes, unless my belief is based on a justified falsehood such as in b2). So, the fact that there are real barns or fake barns that I have not encountered around the real barn that I see and believe that it's a real barn does not bear on my knowledge claim. Furthermore, in cases of direct perception, it seems that in most cases, the reasoning process does not go explicitly as I have shown. That is why it is called direct perception. But I want to point out that even though the reasoning process does not go that way in most cases, that does not mean that we can't think or say that it could go that way. Sometimes, the answers given by someone to the questions about why they believe that something is such and such could reveal that even in cases of direct perception, it can be explained in terms of premise-conclusion.

The fourth counterexample is similar to an example given by Richard Feldman (2003). Feldman calls it *The Extra Reasons Case.* Let's consider the Jones-Smith case again. But this time, the case is a bit different. Suppose I have two independent sets of reasons for thinking that someone

who is in my department will get a raise and he has ten coins in his pocket. This time, I have equally strong reasons for believing that Jones who is in my department will get a raise and he has ten coins in his pocket and for believing that Smith who works in my department will get a raise and he has ten coins in his pocket. But my belief about Jones is false while my belief about Smith is true. Given my reasons, I truly believe that someone who is in my department will get a raise and he has ten coins in his pocket. I form that true belief based on the true proposition 'Smith who is in my department will get a raise and he has ten coins in his pocket'. So, the counterexample goes as follows: It seems that I am justified in my true belief that 'someone who is in my department will get a raise and he has ten coins in his pocket'. But there is an implicit false proposition 'Jones who is in my department will get a raise and he has ten coins in his pocket'. So my belief that 'someone who is in my department will get a raise and he has ten coins in his pocket' is grounded in the false belief 'Jones who is in my department will get a raise and he has ten coins in his pocket'. But we don't want to say that I lack knowledge although some falsehood is amongst my grounds that could lead to the conclusion 'someone who is in my department will get a raise and he has ten coins in his pocket.' So, in this case, although the fourth condition is not satisfied, I know that someone who is in my department will get a raise and he has ten coins in his pocket.

My response to this counterexample relates back to what I have said earlier about the usage of reasons. Suppose that I am sick, and the doctor tells me that I will regain my health in about a week. Suppose that I will really regain my health in about a week and I truly believe that, not based on what the doctor told me, but on the basis that my wife told me so and I always believe what my wife tells me because it makes her feel happy. In this case, I am not justified in believing that I will regain my health because my belief is based on the wrong kind of reasons. But I could have used the information that the doctor gave me, and in that case, I would have been justified. In the *The Extra Reasons Case*, I am not actually using the false proposition 'Jones, who is in my department, will get a raise and he has ten coins in his pocket' as a justification for my true belief 'someone who is in my department will get a raise and he has ten coins in his pocket'. Therefore, my belief that 'someone who is in my department will get a raise and he has ten coins in his pocket' is not grounded in a false belief. The false belief is not used by me as a justification for my conclusion (it justifies my conclusion) and I am still justified in believing that 'Jones who is in my department will get a raise and he has ten coins in his pocket'. If I were to use it, then I would have lacked knowledge.

SECTION 3

Given all those counterexamples, what I want to do in this section is formulate a modified version of the "No False Grounds" account for knowledge to avoid the counterexamples. The original account states the following: S knows P if and only if a) P is true, b) S believes that P, c) S is justified in believing that P, and d) S's belief that P is fully grounded. I will only modify c) and d). We have seen that in any Gettier-style case, we could skip some steps in our reasoning. Just because we sometimes skip some steps in our reasoning, and by doing so, it looks as though our conclusion is not grounded in any false proposition, that does not make it the case that it is not actually grounded in at least one false proposition sometimes. If we consider the *Alternate Route* case, I could have directly concluded 'someone in my department will get a raise and he has ten coins in his pocket' from the information that I have about Jones. It seems that I have skipped a bunch of steps (maybe two steps). But still, the proposition 'Jones will get a raise and he has ten coins in his pocket' is among my reasoning as an implicit assumption and is false. In the *Robot-Sheep* case, I have skipped the proposition 'that animal in the field is a sheep'. That proposition is among my reasoning as an implicit assumption and is false. In the *Fake Barn* case, as Goldman presents it, it seems that there is no explicit or implicit false ground in the reasoning, so Sara satisfies all four conditions. The fact that there are fake barns or real barns does not bear on Sara's knowledge claim. So, Sara knows that what she sees is a barn.

So, on that basis, we want our fourth condition to be something like this: S's belief that P is not based on any falsehood whether it is implicit or explicit. We also want to rule out some justified implicit false propositions from our set of reasoning that do not play a role in the formation of a justified true belief that is not based on any falsehood whether it is implicit or explicit. In *The Extra Reasons Case,* we had to exclude a justified false proposition from the set of reasoning that could lead to the conclusion 'someone in my department will get a raise and he has ten coins in his pocket. 'We want to say that the justified false proposition was not used by me. The true proposition 'Smith who is in my department will get a raise and he has ten coins in his pocket' is the one that was used by me. So, on that basis, I am justified in believing that 'Jones who is in my department will get a raise and he has ten coins in his pocket' but not justified in using it for my conclusion 'someone who is in my department will get a raise and he has ten coins in his pocket'. To say that I justifiably believe a proposition means that I am justified in believing that proposition

and I use that proposition as my reasoning for my conclusion. Thus, we want our third condition to be something like this: S justifiably believes that P. Finally, we could state the "No False Ground" account for knowledge as follows: S knows that P if and only if: a) S believes that P, b) P is true, c) S justifiably believes that P, d) S's belief that P is not based on any falsehood whether it is implicit or explicit.

But there seems to be a worry here regarding the idea of ruling out some implicit propositions that we say are not relevant as reasons that lead to the conclusion. How is it decided whether an implicit proposition is being used or not in the formation of the conclusion? In the *The Extra Reasons Case*, maybe the false belief 'Jones who is in my department will get a raise and he has ten coins in his pocket' is playing a role in the formation of the belief 'someone who is in my department will get a raise and he has ten coins in his pocket. If the implicit false belief is playing a role and the true belief 'Smith, who is in my department, will get a raise and he has ten coins in his pocket 'is also playing a role, then the fourth condition will not be satisfied because there is a justified falsehood involved, but the subject S would still know. We could ask the following: How do implicit propositions get used given that the fact that the Subject S does not reason through them explicitly? Consider this modified example of *The Extra Reasons Case*. Suppose that Fred, who works in my office truly believes 'someone who is in my department will get a raise and he has ten coins in his pocket' on the basis of falsely believing 'Jones who is in my department will get a raise and he has ten coins in his pocket'. Suppose that Fred wants to surprise me and say "Someone who is in our department will get a raise. So, let's go and congratulate him." When Fred says this, I believe that he is referring to Smith. Fred also sees that I got a bunch of presents which I intend to give Smith and congratulate Smith. So when I follow Fred, I say "Why are we heading to Jones's office?" Fred replies, "Well, to congratulate him." I reply, "I already congratulated him this morning. I was about to head to Smith's office and give him those presents." So at that moment, it seems that I was somewhat aware of my implicit false assumption about Jones, but I ruled it out. The implicit false belief about Jones was not being used by me as a justification when I was thinking about the proposition 'someone who is in my department will get a raise and has ten coins in his pocket', although I was somewhat aware of the implicit false belief.

If we consider the counterexamples, it seems that my behavior, expressions, and answers to questions can expose some implicit beliefs that I hold. Even though I don't explicitly reason through those implicit beliefs, still, I am somewhat aware of them. Consider the *Robot-Sheep* case

for example. I was surprised when I knew that there was a sheep with a blue bell tied around its neck. So, my behavior indicated the implicit false belief. I am not saying that behavior will always indicate what sorts of beliefs I hold, whether they are implicit or explicit, but it is a good sign that could point at the implicit assumptions I believe. Even if there are multiple implicit assumptions amongst my reasons, I could still somewhat be aware of which of those implicit assumptions I am considering.

CONCLUSION

What I have tried to show in this paper is that the "No False Ground" account for knowledge is similar to the traditional account for knowledge in holding that the three conditions for knowledge, a) S believes that P, b) P true, c) S is justified in believing that P, are necessary. I have showed that those three conditions are necessary but not jointly sufficient for knowledge. I have shown that the Gettier case challenges the traditional account by showing that a subject S could lack knowledge even though the three conditions are satisfied. For that reason, I had to introduce the "No False Ground" account for knowledge which adds a fourth necessary condition for knowledge. The fourth condition says: d) S's belief that P is fully grounded. I have argued that the "No False Ground" fourth condition along with the other three conditions are jointly sufficient for knowledge. I have also introduced four counterexamples that tries to show that the "No False Ground" is not an adequate account for knowledge. Three of these counterexamples, which are the *Alternate Route* case, the *Robot-Sheep* case, and the *Fake Barn* case, all try to show that a subject S lacks knowledge even though the four conditions are satisfied. On the other hand, the fourth counterexample, which is *The Extra Reasons Case,* tries to show that a subject S knows that P even though the fourth condition is not satisfied. Finally, I have developed a modified version of the "No False Ground" account that avoids those counterexamples.

BIBLIOGRAPHY

Gettier, Edmund. (1963) "Is Justified True Belief Knowledge?" *Analysis* 23

Clark, Michael. (1963) "Knowledge and Grounds. A Comment on Mr. Gettier's Paper," *Analysis* 24, pp. 46-48

Feldman, Richard. (2003) "Modifying The Traditional Analysis of Knowledge," in: *Epistemology*, pp. 25-38 (Prentice-Hall foundations of philosophy series)

Chisholm, Roderick. (1977) *Theory of Knowledge*, 2nd ed, p. 105 (Englewood Cliffs, NJ:Prentice Hall)

Goldman, Alvin. (1976) "Discrimination and Perceptual Knowledge," *The Journal of Philosophy* 73, pp. 772-773

METAWOMAN: A THIRD GENDER IN MENOPAUSE

Colleen Hieber

INTRODUCTION

The rapidly changing perimenopausal body—analogous to the pubescent body—is a battlefield upon which identity is challenged and reimagined. The role of the patriarchal gaze in identity for women during the menopause transition is a salient determinant of social identity through menopause and beyond. In my own experience, my identity has been constructed through a narrative that existed in the spotlight of patriarchy for the last 30 years. As I exit that spotlight, I feel compelled to reflect on the pervasive influence of said gaze on my self-conception and how I might define and reimagine social power that is not contingent on sexual availability or attractiveness once I am out of the spotlight. The patriarchal gaze constitutes women of child-bearing age such that when that gaze is no longer putting pressure on them after menopause, coupled with the biological morphology contributing to a sort of defeminization, she is constituted as not woman, post-woman, or meta-woman—a type of third gender or "ungendered." For this paper, I will examine and put into dialogue the patriarchal gaze, the reproductive body or "proper woman," and Barclay's take on Foucault's panopticism. Through the breakdown of the panopticon for post-menopausal women, I argue that women's claims of invisibility are due to them no longer being socially constructed as women. My use of the term 'woman' is then necessarily narrower than traditionally understood in that I have truncated the length of time a woman can be considered as such for argument's sake. In section one, I will discuss the patriarchal gaze and its effect on gender. In section two, I will utilize Caroline Barclay's take on Foucault's panopticism to emphasize the role of visibility for women during menopause. In section three, I will explore the ways gender is constituted both biologically and socially throughout the major transitions of puberty and menopause. And in the final section, I will argue for the possibilities of a gender beyond womanhood.

This paper will refer to menopause and its various stages several times. In common usage, many women and even their care providers will conflate the term *menopause* with *perimenopause* and *post-menopause*. Since I am discussing identity change across time, the stages and definitions require clarification. *Menopause* is the single day 12 months after a woman's last menstrual period.

Women can hit menopause anytime between 40 and 58, but the average age is 51 in the United States. She may or may not have bleeding after this, but this is the technical term for when a woman is considered to have "entered menopause". It's worth mentioning that this backward-looking metric for menopause is somewhat controversial, and we may see a day when it is redefined (Streicher 2023). *Perimenopause* is the five to ten years leading up to menopause, during which her estrogen levels will erratically fluctuate as they decline, and this fluctuation may cause any or all of the myriad menopausal symptoms (hot flashes, weight gain, change in body shape, brain fog, increased anxiety, depression, insomnia, racing heart, etc.). *Post-menopause* is the stage after menopause, lasts for the rest of a woman's life, and often (but not always) signals the cessation of menopausal symptoms. When a woman is said to be "going through the menopause transition" or "going through menopause", this is a reference to perimenopause. It's important to note that women in post-menopause—where the crux of my argument is concerned—can range in age from 40 to 100.

I use phrases such as *assumed fertility* and *reproductive years*, which include those who are infertile. This is because I am working from the viewpoint outside of her body—people who size up a woman as within a certain age range and make an assumption that she is within the range of ages that women tend to be in to be able to reproduce. Whether the assumption is factually correct is irrelevant to my argument, as I am talking about social judgments, not biological facts.

I will also become entangled in defining various forms of power. The types of power I am referring to are the overlapping varieties of social and sexual—which I am creatively calling social-sexual power—which occur in society at large (I am excluding social environments constituted of entirely or majority post-menopausal women as these spaces are exceptional). This includes the power of being visible to others in social settings, having social influence in sexually charged settings, and being sexually available. It also includes any displays or opportunities for power that extend from environments laden with sexual attributes, e.g., when attractiveness plays a role in getting hired for a job or a promotion, when male desire motivates the inclusion of women in political conversations, or when sexualized bodies of reproductive age are used to sell or influence others in capitalist markets such as social media. There is a default heterosexual lens when discussing male desire due to the subject matter (women in menopause), but I expect that people other than those identifying as women are harmed in this system. In the final section, I will explore other iterations of power, including political, economic, expert, and referent. I will briefly expand

on these to highlight the potential gains of post-menopause. A deeper analysis of these types of power and their contingent nature is worth further exploration but is outside the scope of this paper.

For my consideration of gender, I am working from Judith Butler's theory of gender performativity. Butler argues for the performative nature of gender, whereby gendered behaviors are seen as productive, i.e., a person's behavior produces their gender. Someone who identifies as a woman would need to consistently repeat (or imitate) certain behaviors over time in order to create a stable gender identity. Thus, if these behaviors cease to happen or if different repetitive gendered behaviors emerge, it would follow that a new gender identity is being constructed. I am advocating for a new gender as opposed to an ungendering because of the potential political and social benefits of being at the creation stage of a new gender. In my argument, I redefine the stage of womanhood that occurs during the reproductive years and utilize the term *proper womanhood* to refer to the time before menopause. A necessary part of proper womanhood includes the force of the patriarchal gaze and its expectations of particular gender norms. Proper womanhood is a representation of the set of ideals imposed on women by patriarchy and its values, which include an emphasis on sexual availability. I'll start by setting the groundwork for what I am calling the patriarchal gaze.

SECTION 1: PATRIARCHAL GAZE

I'll begin with how the power of patriarchy works on gender in the social and sexual spheres containing reproductive bodies. The *patriarchal gaze* is a concept that stems from the term *male gaze* as popularized by film theorist Laura Mulvey in her essay "Visual Pleasure and Narrative Cinema" (Mulvey 1999). While Mulvey was describing the subjective male audience that serves to objectify women in film, the patriarchal gaze, as I use it here, represents the visual mechanism of the patriarchal system at large, whereby women are objectified and held to oppressive gender norms by any and all members of society (including other women and capitalist market forces). That is, the patriarchal gaze has the visual capacity to "see" or spotlight that which it intends to exercise power over (women). I mean this to be literal, as done through the act of looking. The power of sight that animates the patriarchal gaze is an objectifying force.

The result of the patriarchal gaze objectifying women's bodies is that women experience their reproductive age bodies (i.e., the female body that is assumed to be capable of reproduction—regardless of fertility status—and, therefore, within an approximate age range of mid-teens to mid-

forties) in a particular way that hinges on temporality. Meaning that women's bodies are in the visual grip of the patriarchal gaze for a finite amount of time, and women are aware of the beginning and end of that timeline. This awareness is oppressive in and of itself. In the years preceding puberty (between 7 and 12 years old (Breehl and Caban 2023)), the female body is not illuminated by the patriarchal gaze in the way that constructs womanhood. The gaze cannot fall on prepubescent girls with the gender-constituting visual force that creates proper womanhood, even through the eyes of pedophiles or sexual deviants, because, by definition, the gender-constituting force on girls creates *girls*—not women. After the reproductive years, in post-menopause, her body is not seen by the patriarchal gaze either. This is not to say that post-menopausal women are not attractive, only that the patriarchal gaze is not exercising its objectifying force on them—particularly the force that constitutes them as proper women. Couldn't other women who have internalized gender norms be constructing post-menopausal women as proper women far into the final years of their lives, thereby nullifying the temporal nature of womanhood? While I don't doubt that gender construction is still occurring for many, if not all, post-menopausal women (for certainly post-menopausal women still act in gendered ways), my position is that without the weight of patriarchy's gender norms related to sexual reproduction and availability, the gender that post-menopausal women are imitating is as different as it was in girlhood. There are exceptions to this, e.g., an 81-year-old Martha Stewart on the cover of Sports Illustrated's May 2023 swimsuit issue is certainly a case of the patriarchal gaze seeing her. But that may be an intractable mixture of fame, wealth, and sex since it was because of her lifetime of success and celebrity that she was asked in the first place—not solely because of her appearance. In both cases of puberty and menopause, the body undergoes visible changes, from an androgynous child to a more curvaceous woman, then following that to post-menopausal morphology (thicker waist, hair thinning, and reduction in breast size and volume). These stages reveal both to the patriarchal gaze and to herself her current status as a former or potential mother and put a priority on sexual availability. The limited nature of a woman's assumed fertility and her perceived vibrancy is consistently culturally reiterated through the ticking biological clock idiom and an internal sense that when her inability to reproduce is foreclosed, she will become social-sexually irrelevant. Her reproductive body is highly valued by the gaze but is ultimately transient.

The transient nature of the reproductive body's value is important because it highlights that whatever force the gaze has on a woman, it is only working on her temporarily, for a minority of

the years of her life should she live into old age. Therefore, if the visual force of the patriarchal gaze is a strong force in the social construction of gender, which I claim it is, such that a woman's gender is contingent on it, then the gender it constructs is also temporary. Additionally, the visibility that the patriarchal gaze temporarily confers to those who are properly adhering to its norms are not just recognized as such but are also bestowed a kind of social power.

Throughout the life of a proper woman, the patriarchal gaze offers her the opportunity to possess or exercise superficial or preliminary power within contexts involving the combination of both social influence and sexual availability. Subsequently, one might capitalize on that opportunity by behaving in ways that attract more recognition—perpetuating feminine stereotypes that communicate sexual availability via the strategic use of makeup, hairstyle, dress, voice modulation, and physical movement. In the same way the male stranger conferred a previously unknown status on me by merely paying attention to me in a particular way, the patriarchal gaze constitutes proper women. This gaze recognizes women who have fulfilled some level of the ageist, sexist, racist, classist, fat-phobic, impossible ideal that patriarchy values. Women are not just morally acceptable when these ideals are achieved (or strived for); they become visible *as* proper women. Once seen as proper women, other people often readily confer the same kinds of social-sexual power to these women that the patriarchal gaze does, for example, conferring attention, respect, a desire to be close, and trust to a proper woman just based on her display of proper womanhood.

Proper womanhood is a term I am using (modeled on Talia Bettcher's *proper appearance* (Bettcher 2007)) to be understood against mere womanhood. Women of any age can display mere womanhood. However, proper womanhood is dictated by patriarchal structures and necessarily includes behaving according to gender norms that prioritize sexual availability, reproductive capacity or youth, and attractiveness. In order to remain visible, a woman must perform gender by displaying proper sexual appearance, which lies within narrow and strict parameters (thin, young, attractive, smiling, etc.). Proper womanhood refers to the idea that a woman's social-sexual power depends on her being seen as an object of sexual desire.

The power that accompanies proper womanhood is contingent on the delusion that an identity based on the trappings of perceived fertility and sexual availability is a liberating power when it is not. It is inherently oppressive since it reduces women's worth to their biology and conformity to male desire. This power also may attract sexual harassment and pursuit, which often

enough leads to violence. This is not to say that women who wield social-sexual power are morally wrong or blameworthy. On the contrary, many women with little to no social-sexual power also get sexually assaulted (elderly, unattractive, larger bodies, poor—all kinds of women get raped). Even without considering the potential for violence, proper womanhood's power links a basic human need for recognition and good social standing to the fleeting timeframe where a woman is capable of reproduction. The bulk of her life—about 40 years—will be after the menopause transition. Thus, there must be a liberatory social power that accounts for those needs.

There is another oppressive feature of identity when it is tied to sexual availability. Mulvaney's male gaze is understood as male desire, effectively coaxing women to display themselves in a way that signals sexual availability and results in male pleasure. However, the patriarchal gaze has an explicit effect on gender. It does more than take pleasure in looking. It puts a demand on women to perform normative gender roles. This necessarily negatively affects men, non-binary, and trans people, as well. Punishment for deviation is the threat of invisibility (or much worse), which is tantamount to being cut off from power. This disciplinary force recalls Foucault's panopticism.

SECTION 2: BARCLAY'S PANOPTICON BREAKDOWN IN THE MENOPAUSE TRANSITION

In an attempt to capture the implication of the subjective experience of invisibility for women in menopause, I will look at Caroline Barkley's application of Foucault's panopticism to gender norms. Foucault analyzed Jeremy Bentham's architectural idea of a panopticon, which is a central tower surrounded by a circular building whereby those in the tower can see those in the surrounding building, but the people in the building cannot see those in the tower. The important element is that those in the surrounding building (imagined as inmates, mental patients, workers, or students) are certain that they could be being watched at any time. When Foucault applied the panopticon to society as a social theory, each person was the central observation tower. Therefore, "All individuals are then both the watchers and the watched" (Barkley n.d., p. 10). Discipline is passive, as people readily police their own and others' behavior. In Foucault's analysis, visibility is the key component, "Hence the major effect of the Panopticon: to induce in the inmate a state of conscious and permanent visibility that assures the automatic functioning of power" (Foucault 2008, p. 6).

Caroline Barkley also reiterates the importance of visibility, "This constant visibility is an essential facet of panopticism in which society prevents the individual … from acting outside of the social expectations" (Barkley n.d., p. 11). Acting in appropriately gendered ways (presenting traditionally feminine aesthetics) is a requirement to avoid punishment. The gender norms and mechanisms of the social panopticon are not threatening to those who have internalized them, though, which creates an intractable system (Barkley, n.d.). Additionally, the diet or wellness, fashion, entertainment, marketing, and beauty industries, which rely heavily on a sexist, ageist aesthetic, benefit from such a structure as older women are constantly bombarded with messaging that encourages them to spend time and money trying to stay visible. Being visible in the panoptic structures can even seem enjoyable if adherence to the norms comes easily, i.e., if you are young, thin, light-skinned, attractive, etc. However, what does this mean when post-menopausal women complain about not being seen or being invisible?

In a real sense, women are controlled by panopticism, even while alone, made to constantly strive to prove that they are women throughout their reproductive years. Once past menopause, many women continue to vie for visibility; that is, the patriarchal panopticon wields power over those it does not even see. Eventually, women might realize they are not visible to the gaze—which would equate to feeling irrelevant to the system of social-sexual power they were once a part of. It is likely that feelings of irrelevance are felt more strongly—and maybe only—by those women who are highly compliant with its gender-constructing force, i.e., those who managed to closely approximate the ideals of proper womanhood and benefit from its inherent social-sexual power.

Continued attempts to reengage with the gaze (e.g., through cosmetic surgery, anti-aging products and procedures, the continued use of makeup, wearing high heels, etc.) might feel like the only way to claim womanhood—which has become a key signifier of their identity. However, recognition, connection, and social influence can be accommodated in other ways besides conforming to the patriarchal gaze if a critical mass of women prioritizes creating a new system of values (or amplifying one that already exists). Not only might this demotivate the stubborn attempts to reengage with the patriarchal gaze, but all marginalized people would benefit from a liberatory system of visibility that can be sustained throughout a lifetime.

Two questions arise here:

(a) If there is no panoptic force granting and affirming gender—do you have one? That is, if there is no panoptic observer to display your gender to (because they have stopped looking, and you know it), are you ungendered?

(b) Is there a third-gender option?

To answer (a), we have to consider that those visible to the panoptic patriarchy are certainly held to gender norms. Those who distance themselves from these norms, non-binary people, for example, are still negotiating the disciplinary force of the panopticon. They actively resist it. To do that, they often must identify strongly with an alternative gender identity. However, they do not disregard a position on gender outright. Even so-called gender-fluid and agender people are situating themselves against the matrix of a gender identity spectrum—if only to firmly claim to be at the far end or outside. Additionally, the majority of the social sphere is policed into acting in alignment with gender norms. Cis people necessarily understand their identity as being contingent on the identities they impute to others around them. For example, if a person sees men, women, non-binary, and trans people, they will likely recognize themselves in some gender category. Although, for post-menopausal women, the gender category that resonates the most may be one that hasn't been mentioned yet—a gender that reflects their unique combination of biological changes and changes in social-sexual power.

While it may be true for some, I do not think ungendering is necessarily happening for post-menopausal women due to the fact that we repeat performative gender acts even while alone or being ignored. Part of panopticism is the expectation that we could be being watched at any time—even if, in reality, the people in the central tower are not looking (a fact that may be true for post-menopausal women)—which creates a sort of hallucination of expected gender norms that nevertheless produces gendered behaviors. Ungendering could be an intermediary experience between proper womanhood and whatever comes after, though.

To answer (b), I will start by pushing against the claim that middle-aged women are fundamentally invisible. For a social object to be invisible, there must be an initial moment of visibility, then a swift "binary division and branding" (Foucault 2008, p. 4) of interpersonal objects as sexually relevant or not sexually relevant, and finally, a judgment of visible or invisible. So, it is incorrect to claim invisibility; rather, women at this life stage are designated as outside the

disciplinary power of the patriarchal panopticon because they are visible as something other than a proper woman. The visual force of panoptic patriarchy sees the post-menopausal women, then judges her to be in violation of its ideals. The ungovernable, no longer proper, post-menopausal woman has transitioned into what I will call a *metawoman*. She is not punished, though; she is now free to claim a new identity.

Metawoman can be seen as a new gender due to the change in social pressures that are put on women as they transition out of their reproductive years. They are no longer expected to care for children in the same way, their bodies are less objectified, and they are less likely to be visible in social-sexual settings. The biological and social change can also be ritualized in order to put the menopause transition in a positive light.

I believe creating a gendered distinction for women in menopause will serve the political and social needs of this group of women. Creating an identity-based politic for this group could serve to improve areas of their life where they continue to suffer from age and gender-based discrimination. For example, it is common for medical studies to exclude women in menopause altogether or not make note of their menopausal status. In the flawed 2002 Women's Health Initiative study, they were lumped together with elderly women in their 80's, which resulted in menopausal hormone therapy (MHT) being universally condemned for all post-menopausal women when in fact, it is only problematic for women who *begin* taking it more than 10 years after menopause (Cagnacci and Venier 2019). This caused most women and the healthcare world to disavow MHT as dangerous, with its risks outweighing the benefits, thereby depriving women of needed healthcare that would have greatly improved their quality of life. Although the reevaluation of that flawed study was followed by further studies (over the course of 13 years) showing the health benefits of correctly timed MHT, those new findings did not receive adequate media attention, and the unfounded fear of MHT continues for women going through menopause. A politically motivated gender identity centered on post-menopausal women could advocate publicly for these types of injustices. Social visibility must include being accounted for in the data.

Rather than be distracted by vain and confused attempts to recapture the oppressive gaze of patriarchal panopticism, a metawoman is a liberatory position. One that women can look forward to. I am under no illusion that post-menopausal women will embrace a gender that disavows the identifier "woman". In search of a descriptor for a third gender option, I found "meta-" is defined as "denoting a change of position or condition" and "denoting something of a

higher or second-order kind". I have no prescriptions for what normative gendered acts are applicable to metawomen, if any, but the experience of other transgender people can offer insight.

Darcey Steinke appeals to the experience of trans people navigating between the sexes to understand her own experience of ungendering through the menopause transition. She describes her femininity as "fraying" and willfully sits in the experience of being twice misgendered, called "Sir", and escaping what she describes as a "claustrophobic femininity" (Steinke 2019). Transman Max Wolf Valerio explained his transition as an "erotically charged boundary crossing". He explicitly compares his mental experience of transition to menopause and describes changing [sex?] as an investigation, an adventurous opportunity "to live beyond the given and the commonplace" (Valerio 2006). Julia Serrano describes her transition as "an ungendering, an unraveling" and highlights the close proximity of men and women as "only a hormone prescription away" from each other (Serano 2016). Steinke concludes her exposition of menopause by asking whether each of us should be pursuing a synthesis of the sexes. These are examples of the way gender change can be seen as positive and therefore embraced.

As a historical case, in Miranda's essay "Extermination of the *Joyas*," the third-gender *joyas* in California Indian communities (who presented similarly to transwomen today) held specific responsibilities, including ceremonial rituals around death. Their gender roles were analogous to those of the women of the tribe but also went beyond them. Similarly to the patriarchal panopticon, when Spanish colonizers, including the Catholic priesthood, observed the *joyas*, they were not constituted as women. Interestingly, *joyas* and post-menopausal women were peers, "The threshold of death was the realm of the 'aqi, and no California Indian community was safe or complete without that mediator. Asserting [sic] that undertakers were exclusively 'aqi or post-menopausal women (also called 'aqi)" (Miranda 2010, p. 266). Although Miranda does claim that a third-gender status seems to have been extended to post-menopausal women, she then, unfortunately, conflates "elderly" women with post-menopausal women. To clarify, Michelle Obama or Jennifer Lopez would not be considered "elderly" even though they are older than the average age of menopause. The joyas offer us an example of what menopausal social power might look like. This also suggests there may be a fourth gender extended to women who would identify as elderly or simply would not identify themselves in terms of their menstruation status if it has been decades since its cessation.

SECTION 3: CONFIRMATION OF GENDER FROM GIRLHOOD TO WOMANHOOD AND BEYOND

To understand post-menopausal gender imputation, I will analyze the analogous process of gender imputation during the transition from girlhood to womanhood. I believe there is little to no metaphysical difference between the transition from girl to woman and the transition from reproductive-age woman to post-menopausal woman. I would argue that girls and women are not precisely the same gender, especially because "gender more often refers to cultural and social differences" (New Oxford American Dictionary). The emphasis on social and cultural contexts in the definition supports my position since social invisibility during menopause is the claim. Both transitions last several years and involve a change in various metaphysical properties, including identity, relationality, agency, and embodiment, not to mention biological changes. The coupling of perimenopause with entering mature adulthood might put more pressure on the menopause transition in the sense that there is more at stake for the metaphysical changes. For example, the practical sexual freedom inherent in the cessation of menses and the embodied change that reduces the need for women to make themselves sexually available or vie for male desire could have farther-reaching effects for an independent person as opposed to a child. It seems that after puberty, a girl or young woman is far more socially constrained in terms of her behavior and society's expectations of her than after menopause.

The biological component of the transition from girl to woman is often determined in relation to reproduction and, thus, sexual availability. Colloquially, a girl *becomes a woman* once she is apparently physically able to reproduce and raise children, a comment many young women hear from their families and in media when she starts her period. The shift is often made known by a change in body morphology, the onset of menstruation, breast development, and hormonal changes. This cluster of changes temporally coincides with society beginning to confer the status of womanhood to a person previously regarded as a girl or child—a process that takes several years. What happens sooner than womanhood is the recognition of her sexual availability. The coupling of internal bodily change and societal recognition of sexual availability constitutes an identity that connects the inner life and the outer life. That is, a girl becomes a woman throughout the time when internal biological changes are happening, and her gender is edified by newfound social-sexual power.

When I was a young teenager, I explicitly remember interacting with an adult man in a way that represented the first time I was no longer regarded as a girl, the first time that I would begin

reiterating behaviors of my new woman gender. Before that moment, adult men regarded me as a female child (if I was regarded at all). There was nothing remarkable about being treated specifically as a child, but the shift to being treated as something other than a child is worth remarking on. The interaction was brief and trivial—I had a situation that a stranger helped me with, the details of which I have since forgotten. It was memorable not because it was a flirtation per se but rather because the man took me seriously with his posture, time, and communication. I seemed able to command his attention. It was memorable because I realized at that moment that the world regarded me differently than it had previously. I had been gifted a sort of social power via social recognition. From that moment forward, I was aware of and often utilized that social power whenever it was available. At the time, it did not seem that this was a power I created or found within myself. It appeared to be a power that was granted to me by others—men and women—and exploited by capitalist markets such as the beauty and fashion industries. This was not simply impending adulthood but maturation infused with sexual availability. I think this because the interactions had an intimate and playful nature to them. Additionally, being aware that I had been given some clout fueled my confidence to command that social power repetitively, effectively doubling down on the fact that my visibility was tied to my sexual availability. I repeated behaviors that served to constitute myself as an agent with social standing (à la Butler) via my newly assigned gender as a woman.

By both repeating and receiving social power from others via gender imputation (as a woman), I repeatedly experienced an external confirmation of my gender. This was accompanied by an internal sense that I was no longer a girl and was on my way to referring to myself as a woman—although this would be a relatively lengthy transition. I would not claim the title of woman until sometime in my late twenties, and I would do so timidly. But why timidly? Perhaps because society would continue to refer to me as a girl in a litany of situations, personal and professional. Or because I was often sexualizing myself as the youngest version of an adult. Maybe because "girl" allowed me to abdicate responsibility. Regardless, I felt the title of "woman" to be too big a claim, an uneasy self-reference. Calling myself a woman felt insincere when I was 23 and still strange at 27, well into adulthood. The lag between how I was viewed by others and how I viewed myself - an adult menstruating girl masquerading as a woman—is repeated in reverse during the menopause transition. I am now a mostly non-menstruating adult female, questioning in what ways the woman I've historically been is expected to interact with others. The way I

conceive of my own gender is starting to break down as I am often not seen by the forces that impose gender norms, and I feel my social-sexual power diminishing. Since my internal biology has changed significantly with the radical drop in estrogen, my diminishing social-sexual visibility is not strong enough to constitute my proper womanhood, and social visibility remains a psychological and political need or desire; in the final section, I will introduce some new possibilities for women who want to be visible as metawomen.

SECTION 4: A LIBERATORY FRAMEWORK FOR METAWOMANHOOD

There are benefits in escaping the patriarchal gaze after menopause, including more freedom from caring for young children (although this care requirement is often redirected towards aging parents), not feeling compelled to strive for unrealistic beauty standards, and being less likely to be a target for sexual harassment; not to mention no longer being beholden to the biological ups and downs of a menstrual cycle or fears of unwanted pregnancies. But that is merely what you don't get. A truly liberatory position must offer some kind of self-determination that capitalizes on the wisdom of experience that older women have. Since the major social changes of menopause include the loss of social-sexual power, it is important to highlight other types of power that can be readily utilized by metawomen.

The kind of power afforded to metawomen might be simply the power inherent in being seen, which offers an emotional salve and can boost confidence and self-esteem. Being able to be seen is a sort of soft power that allows for a humane and respectful connection with others. As other forms of power naturally have an expiration date (e.g., positional power that is tied to a job title), this kind of power can persist throughout life. Importantly, this reinstates the visibility that was present in panoptic patriarchy without the oppressive and sometimes violent gender norms associated with proper womanhood. It's worth noting that many older women find themselves in a better financial position when they are older, thus increasing their economic power. In a world where metawoman is visually recognized as such, her position at work, in political spaces, and even in athletic competition would be elevated. Referent power, or the type of power that a mentor or leader has, is a political power that can be claimed. In academic, corporate, or sports settings, expert power can be harnessed, especially for women who have maintained long careers in the same field. While these powers are currently exercised by women, if communities of metawomen

placed an emphasis on them (and men supported them), younger men and women might be encouraged to seek out metawomen for their particular skills.

If metawomanhood were seen and respected as a time in life when power born of experience was at its height, perhaps gender norms for metawomanhood would fall away. I have often heard women in their fifties talk about being free from caring what others think, which, ironically, is sexy. So while gender norms for metawomen would still involve sexual attractiveness, they also could be centered on pleasure and partnership rather than procreation and motherhood. These forms of power (unattached to sex or bodies) combined with a social-sexual existence that does not prioritize motherhood might be exactly the way many women would love to age—with a strong sense of self, born of experience and wisdom, expressed through a gender that was created by and for them.

That metawoman is a third gender, discoverable post-menopause, is made clear by understanding how women are socially constructed. By understanding the way gender is policed through panoptic patriarchy and how that structure relies heavily on the presumption that we are visible, we can see that women transitioning through menopause can harness the subjective invisibility they experience in order to claim a new liberatory gender. In the end, we do not need to wait for patriarchy itself to be dismantled before recognizing that becoming a metawoman after menopause is a liberating rite of passage.

Bibliography

Barkley, Caroline. n.d. “Panoptic Patriarchy and the Gap between Queer Lives,” Accessed December 4, 2023, https://onyxreview.agnesscott.org/wp-content/uploads/2016/12/Barkley.C_PanopticPartriarchy.pdf

Bettcher, Talia Mae. (2007) “Evil Deceivers and Make-Believers: On Transphobic Violence and the Politics of Illusion,” *Hypatia* 22 (3), pp. 43–65

Breehl, Logen, & Omar Caban. (2023) *Physiology, Puberty,* (StatPearls Publishing)

Cagnacci, Angelo, and Martina Venier. (2019) “The Controversial History of Hormone Replacement Therapy,” *Medicina* 55 (9). https://doi.org/10.3390/medicina55090602

Foucault, Michel. (2008) “‘Panopticism’ from ‘Discipline & Punish: The Birth of the Prison,’” *Race/Ethnicity: Multidisciplinary Global Contexts* 2 (1), pp. 1–12

Miranda, Deborah A. (2010) “EXTERMINATION OF THE *JOYAS*,” *Glq* 16 (1-2), pp. 253–284.

Mulvey, Laura. (1999) “VISUAL PLEASURE AND NARRATIVE CINEMA,” in: S. Thornham (Ed), *Feminist Film Theory: A Reader*, pp. 58–69 (Edinburgh University Press)

Serano, Julia. (2016) *Whipping Girl: A Transsexual Woman on Sexism and the Scapegoating of Femininity* (Basic Books)

Steinke, Darcey. (2019) “Going Through Menopause Changed The Way I Think About Gender,” BuzzFeed News, June 18, 2019, https://www.buzzfeednews.com/article/darceysteinke/i-felt-confined-by-femininity-for-most-of-my-life-menopause

Streicher, Lauren. (2023) “Is It Time to Rethink How Menopause Is Defined?” *Menopause* 30 (12), pp. 1177–1178

Valerio, Max Wolf. (2006) *The Testosterone Files: My Hormonal and Social Transformation from Female to Male* (Basic Books)

THE IMMORTAL LIFE: ITS DESIRABILITY AND CONSEQUENCES

Cyprus Marques

INTRODUCTION

Immortality as a subject of discussion involves several intriguing philosophical considerations. Following technological developments in our lifetime increasing the longevity of life for humans like never before, it is worth asking ourselves moral questions regarding the desirability of immortality should humankind ever grasp it in their mortal palms. The main purpose of this paper will be to show that many arguments against the desirability of an immortal life for the individual are unfounded. In Section I of this paper, I will argue that Bernard Williams's identity assertions regarding the desirability of an immortal life are too strict and result in some counterintuitive notions. In Section II, I consider the consequences of an immortal life, and I challenge suppositions regarding the inevitable boredom of such a life by showing it would be quite fulfilling within certain constraints. Then, I will argue that a loss of human values such as urgency, courage, etc. frequently posited is not an inevitable feature in an immortal life. Finally, in Section III I argue that while a corporeal immortal life of fixed abilities is still desirable, the most exciting form of immortality available to us would be one in which our Earthly limitations have become surpassed by our technological discoveries and developments and our abilities are no longer fixed.

Before we can begin a thorough investigation of the desirableness of immortality, we ought to clarify what it is we exactly mean by the term when we postulate it. As it stands there are several forms for us to explore, each with its particular implications. However, for the sake of this paper, we will focus mainly on *corporeal immortality*, or eternal life as the same individual in the same body infinitely extended to the future. There will be one exception on this point we will discuss, as this form of immortality may eventually involve the discarding of the human body in order to take a different corporeal form as technological developments become more sophisticated, as we will see might be possible in immortality with non-fixed abilities.

Secondly, I will be operating from the presumption that upon death, our consciousness, personality, and unique psychologies, as well as any experience, memories, selfhood, and the like derived from that will completely cease to be. I will take it for granted that there is no afterlife or change in existence after death, just the unspeakable void upon which everything we are ceases to

exist. It is from this assumption that the desirableness of immortality will be the most usefully assessed, as on occasion we will be comparing the potentials of immortality with the threat of nonexistence.

SECTION I: IDENTITY CONSIDERATIONS

Bernard Williams (1973) offers a method of assessing immortality and its desirability. Williams postulates that two general camps of desire are relevant here: *Conditional desires,* or desires that do not provide us with a reason to continue living, contingent based on environment, for example, desire for water, food, shelter, etc.; and *Categorical desires,* which provide us with a reason to desire to continue in life, such as raising happy children, having an effect on one's community, and so on. The distinction, Williams argues, is important because conditional desires provide us with more sophisticated long-term reasons to want to continue in existence.

Williams asserts that there must be a relevant connection between the categorical desires of my future self and that of my current self for immortality to be desirable:

> …one of two important conditions which must be satisfied by anything which is to be adequate as a fulfillment of my anti-Lucretian hope, namely that it should clearly be *me* who lives forever. The second important condition is that the state in which I survive should be one which, to me looking forward, will be adequately related, in the life it presents, to those aims which I now have in wanting to survive at all (Williams 1973, p. 9).

He insists on these conditions because, for one to consider a future life desirable, the future self must in some way respect their desires before the fact. In considering the possibilities for immortality, a person could not have any hope that their future life would contain any resemblance to their current life, which would mean that there would be nothing by which for him to judge the desirability of it provided extreme changes to his character occurred. Here I will push back on Williams by introducing what I see as a few flaws in this manner of thinking, as well as the strictness of the postulates that he has asserted.

Firstly, constructing identity this way seems entirely too restrictive in terms of delineating identity and desire. For one, humans often evolve and change in unforeseen ways, developing new desires, new traits, and characteristics based on their environment and their interaction with it,

some of which may have not ever been anticipated by a past self. Christine Overall has also pressed Williams on this point equally. She states:

> … his conclusion that the failure to meet these conditions creates an insuperable barrier to the desirability of immortality reflects a lack of imagination about the range of life's possibilities … I would guess that most people can imagine several lives, quite different from their present ones, that they would have greatly enjoyed and flourished in if they had had many more years in which to develop them and the opportunity to choose them (Overall 2003, Sp.159).

And following this, our developed self is capable of looking back on the old self's desires and still recognize it as being continuous with their own identity; I doubt an immortal life would instantaneously involve extreme personality changes such that one becomes a drastically different person every 100 years, say. What is more likely to occur is a continuous sort of stream of personality changes which happen very slowly such that they aren't noticed until one looks back at what has changed. And, because of the nature of human personalities, it's equally possible that there will still be some core character traits which maintain themselves anyway, though certainly not all will be immune to change. If one has a particular sort of personality in which the fulfillment of desires could still be maintained over a continuous existence, regardless of the specific character of their desires, it seems fair to say that desiring immortality is at least somewhat a rational possible desire to have on the same basis.

Similarly, Altshuler (2015) discusses that the sort of feelings that arise when considering immortality and change of character depends on one's dispositional manner of looking at the possibilities. *Forward-looking,* or looking from the perspective of a self now with the expectation that it will someday change, results in wariness regarding the future self ultimately swallowing up the old and having the old self's desires becoming obsolete. *Backward-looking,* however, results in a "wait and see" attitude as one can see the benefits of having changed or is pleased with their current state and as such experiences no panic regarding the matter. Altshuler states, in my opinion wrongly, that the forward-looking perspective is indeed the best manner of considering the possibilities, even though both methods have their respective virtues. In as much as we are discussing the potential of an immortal future, we must look ahead as opposed to looking

backward, because it is only from the present self's perspective that considering the desirability of a future makes any sense. We cannot look back on ourselves ahead of time or predict in any manner what exactly will occur in the far future, so we ought to stick to the forward-looking perspective in order to judge anything.

I believe that the forward-looking perspective is misguided in considering immortality, in that it renders the nature of the judgment dependent on a limited self's purview. We must indeed engage with the current self to even determine the desirability of a future immortal state at all, but it is also fair to say that the backward-looking perspective offers us a significant reason to consider why any anxieties we might have would be completely unfounded. Curiosity and openness are prerequisites here to avoiding the anxiety related to identity changes. It is true however, as Altshuler states, that the desirability of such a state would still be under question, as we cannot desire a future in which the circumstances might no longer be relevant or fulfilling to the current self we are. But I think Altshuler, and Williams equally, place too much weight on the specifics of desires as opposed to concentrating on the experience individuals wish to have as a byproduct in a lifetime. One has to be open to the idea of different desires and different future perspectives potentially affording us the sort of satisfaction and occupation we seek in our own mortal lives. It affords us a level of peace with the idea that our desires are ultimately not permanent portions of our selfhood, which without the self inevitably dissolves and becomes irrelevant. As such, we can then have a rational discourse about the desirability of immortality on the same basis, assuming from a wiser point of view that while we might not share the same desires as our future self, that we will still be the same individual, and still retain our access to pleasurable and desirable experiences regardless of their specific character.

SECTION II: AN ASSESSMENT OF CORPOREAL IMMORTALITY AND ITS POTENTIAL CONSEQUENCES

For the purposes of this paper, we are going to limit our discussion of immortality to a particular type, namely one in which psychological continuity is maintained, and the body maintains its health and wellbeing indefinitely. For some other useful preconditions, Christine Overall (2003, pp. 128-131) provides useful stipulations for how this immortality ought to function, some of which I will mention because they are relevant to the conversation. Firstly, individuals would be able to choose to be immortal, and secondly and importantly, they could choose to cease being

immortal at any point and are not forced into continued existence should they wish to die. Thirdly, individuals would essentially maintain good and vibrant health throughout the immortal period. Disabled peoples, people born with health issues or chronic illnesses, and so on, will presumably at this point in time would ideally have access to some form of a cure or workaround which will allow them to escape their compromised state before they become immortal, and so this form of immortality would be open to quite literally everyone.

Now, I wish to focus the majority of this section on whether or not immortality of this form is indeed desirable for the individual, as I believe there are more interesting arguments to be deliberated over from an individual perspective as opposed to engaging in some of the larger social dilemmas which result from the immortality of the masses. Suffice it to say in the model I am considering, there will likely be other immortals available, but being immortal as a human is not an absolute rule, there will likely be individuals who choose to pass into death naturally as we do today. With that caveat forward, we can move on to consider why many have considered immortality largely unsavory.

2.1: Boredom

Williams in his paper on the subject proclaims that the nature of an immortal life is one in which there will inevitably be a sort of unrelenting and morose sense of boredom permeating one's life experiences past a certain point. He posits that for us to desire immortality the individual in question ought to have some sort of fixed character, based on the suppositions I have addressed earlier in the paper. If this is so, it seems inevitable that a fixed character existing on Earth will eventually undergo all the possible experiences available to them from their categorical desires, losing their wonder for the world with their life eventually becoming unsavory and predictable. Such a lack of novelty, Williams argues, would be unbearable, and this means that from our current perspective, an immortal life must be hopelessly banal and thus undesirable.

We have the strictness of Williams's identity claims, as I have tried to deconstruct above, to contend with here. If one is fixed along the same lines and rigid in their character then I have no doubt they will continue to be bored; but as I have argued above, it does not follow that an immortal life in which one's character will change drastically or continuously is not desirable, especially considering many humans are indeed capable of desiring changes in personality. But this is not the most interesting counterpoint to Williams's claims, I think.

There are some things we have to consider concerning how Williams justifies his argument. He relies heavily on a book character named Elina Makropulos, a woman who has consumed an immortal elixir and lives to the age of 342. He says of her characterization in the story: "Her unending life has come to a state of boredom, indifference and coldness. Everything is joyless: 'in the end it is the same', she says, 'singing and silence'" (Williams 1973). It might seem at first that one ought to read this analysis with skepticism seeing as how Williams relies on a book character to make his claims. After all, it seems hardly clear to me that a book character represents an objective or empirical study of the possibilities underlying an immortal life; perhaps, someone might protest all the same, it is our only viable option considering immortality of the sorts in question is a fantasy anyway, so there cannot be much more we can use other than our imagination to determine its desirability. But I think this represents a failure of imagination, precisely because their views aren't sufficiently informed by empirical and psychological considerations relating to the possibilities of immortality. This is not to say that fiction novels cannot offer an emotional account of how some experiences might go, but that alone they cannot be taken at face value in settling an argument on the manner.

Bortolotti and Nagasawa (2009) argue along these lines, stating that Williams's arguments lend too much to the conceptual when research has now been conducted which renders some of these conceptual suppositions under scrutiny. Of course, Williams was not around when this data was released, but modern developments suggest a more complicated picture than he puts forth. Bortolotti and Nagasawa provide a review of psychological research on the nature of boredom to illustrate some of these new discoveries. They express the difference between two different types of boredom: *Situational boredom*, or boredom characterized by low levels of arousal in response to under-stimulating situations, and *habitual boredom*, which has been found to be somewhat of a psychological character trait. The second type of boredom is a direct byproduct of one's psychology and *not* a result of a reaction to one's life; if the research conducted holds any scrutiny here, it seems fairly easy to argue that if Makropulos finds herself so devastatingly bored then perhaps what she suffers from is not situational boredom, which is environment oriented, but from habitual boredom, in which her character or psychology in some way suffers an affliction which causes her to tire greatly of life.

If it's true that there is a distinction between the two types of boredom, then the challenge seems to lie not in relieving habitual boredom, which in any case would probably require *more*

than just new and stimulating activity to resolve, but rather in managing situational boredom as it stands. This seems incredibly easy for one to do; we live in a world in which there is nearly an infinite variety of possibilities and experiences available. And, as it stands, we already do this in mortal existence anyway. Few of us have the luxury of constantly being busy and occupied with activities that interest us deeply, and as a result we attempt to find activities to fill the space and time we have available with things that engage us. Thus, it is unlikely to be any different to an immortal, who will have to discover new activities by which to keep themselves engaged, satisfied, and occupied. When one's time is literally limitless, the very framework by which one can analyze and create new projects is opened up so to speak to some really interesting possibilities; suddenly activities that would have previously taken many lifetimes to achieve, activities which under our current life span would hopelessly be cut short, are now open to us for development.

The nature of our current existence often prompts us to be modest, if we are honest with ourselves, about our capabilities to achieve given our limitations. An immortal, however, can look upon certain temporally strained tasks with wild abandon and hope. Overall has something to say along these lines as well: "Another example of an unending and boredom-defying activity is the quest for wisdom or enlightenment. In many traditions, the fulfillment of that quest is thought to require many lifetimes. An immortal life on earth appears to provide the opportunity for taking seriously the pursuit of satori" (Overall 2003, p. 151). If we can imagine that there would be projects that could reasonably take an incredibly long time to achieve, then we can take it for granted that immortal beings could occupy themselves with them and thus avoid the boredom Williams so insists is inevitable. It might be granted that being occupied does not necessarily mean we will be *entertained* by such activities, even as long standing as they are. But considering the infinite expanse of time available, one could easily spend a few years trying to find specific long-term goals and hobbies they could engage in deeply and meaningfully.

Another line of argument by Fischer (1994) in his response to Williams's claims asserts that there are two different types of pleasures at question here: *Self exhausting pleasures*, or pleasures which, once accomplished, will not result in the desire to repeat it, for example, one's wedding day, or for some women giving birth. Then there are *repeatable pleasures* or pleasures which, provided they are managed correctly and spaced apart sufficiently, will always need satisfying (such as a desire for food, sex, companionship, and so on). Fischer admits that it is true that self-exhausting pleasures would result in boredom if they were the sole source of a person's

interest in a given lifetime, but that repeatable pleasures are a rational and considerable source of entertainment from which to pull. It might be too far to suggest that repeatable pleasures alone would be sufficient to scare off the boredom Williams is concerned about, but this does not have to be a fear if one has other long-term projects and interests like the ones I have discussed above. Rather it seems obvious that an immortal, like many humans already do, would need to learn when to manage boredom by ensuring they have a plethora of options and entertainment from which to choose at any given time to stave it off. They could not rely on one thing alone, however.

There is here one point which I will concede, and that is that those without ambition or interest in life as it stands will likely find the prospect of immortality unbearable. If one has no interest other than staying alive and has no categorical desires or hopes beyond it, then indeed they might find themselves in the place of Makropulos, extending their time indefinitely when they have nothing left to pursue or interest in doing so. So I will concede that those without any ambition or lust for experience or life should disavow an immortal existence; I have no reason however to accept that a forward-thinking and imaginative person would have the same troubles as if we take it for granted that such an individual must have a fixed character, their ambitiousness will likely follow them through the rest of eternity, keeping them busy with new projects, new desires, new means of expression and experience.

There is one intriguing line of argument however which I remain agnostic on, and that has to do with the repetitive character of human relationships; I can imagine an individual might find themselves perhaps not bored but frustrated with the personalities of the mortal humans around them. There are only so many times you can watch a human make the same mistake over and over, especially over various lifetimes, before one becomes sickened with humankind and their behavior and the fact that they never seem to learn from their mistakes. But there are two points against this which I think stand well: 1. Granting that there are other immortals present, we can assume that these personalities would inevitably learn from these repeated conflicts and learn to manage them in satisfactory ways, surpassing the mortals in their psychological development and eschewing the repetition of the same problems. Thus, one's immortal peers can provide a source of novelty and development to pull from in terms of their relationships. 2. This development could feasibly result in a sort of impressive patience and compassion, if an individual deliberately seeks such practices out and is disciplined in their application of them. It might seem hard to imagine that any human would obtain such guru-like serenity in the face of sheer human stupidity or folly, but it certainly

is a possibility for an immortal with all the time of the world in their hands. I say I remain agnostic on these points because we cannot be certain that such things will follow from an immortal life in the same way as our pursuit of satisfying desires and activities, which seem feasible even from our own mortal perspectives. But they do seem significant enough as counterpoints that I feel as though they prevent any serious concerns from arising for an immortal existence.

Keeping that in mind, let's move to a more thorough analysis of how such changes might be said to impact our values as humans. After all, there is more to life than the mere satisfaction of desires and activities, so what of the human desire for meaning and values?

2.2: Human Values and Death

Some individuals proclaim that immortality would result in a significant shift away from deeply human values which provide us with a much-needed sense of meaning to our interactions and projects. In the words of Martha Nussbaum: "Our finitude, and in particular our mortality, which is a particularly central case of our finitude, and which conditions all our other awareness of limit, is a constitutive factor in all valuable things' having for us the value that in fact they have" (Nussbaum 1989). The arguments for such a view tend to go as follows: The temporary, finite nature of our existence places us in a precarious position as it relates to our projects. The fact that we are limited as humans, that we are aware of our vulnerability and mortality, adds various values to our considerations which would not follow in an immortal life. Take courage, for example: Without the vulnerability of death, an immortal cannot be courageous because there is no true risk involved in their behaviors. Nussbaum says, rather curiously, that the immortal gods of the Greek pantheon would not have been considered to be courageous, because there is no manner in which their actions would put their life on the line.

I find this argument puzzling for two reasons: Firstly, courage consists in far more than being willing to put oneself at risk of death in trying to achieve some end goal. This is a very, very narrow definition of courage which I think would severely limit its reading in acts we would typically consider quite brave in nature. Take, for example, a whistleblower working at bank that helps their wealthy clients avoid taxes. Now, they know that should they reveal this information to relevant law enforcement personnel. They are unlikely to be murdered, as their superiors and those paying them don't have any interest in killing anybody. But they do know that they stand to lose their well-paying job, subject themselves to criminal proceedings and interrogation,

potentially be black balled in their career and essentially have any of their possibilities for a successful financial future destroyed at the hands of the powers that be. Would their decision to still go to law enforcement, despite knowing this is a serious possibility, not be considered courageous merely because death was not an immediate possibility? Courage, it seems to me, centers around taking action and doing the right thing even if someone stands to lose something or suffer significantly. That can include someone potentially fighting despite knowing they might stand to die, but it also might include someone who knows speaking out would cost them a great deal of their fortunate circumstances.

Similarly, under the model I have posited, the experience of extreme pain is still not out of the picture for an immortal either. To understand this, take the myth of Prometheus, the god who chose to reveal fire to humankind against the strong interests of his superiors (Hesiod 1983). Prometheus likely understood that there would be a severe punishment as a consequence. His punishment, in the form of an eagle that repeatedly and without exhausting eats his regrowing liver each and every day, stuck in eternal torment, presents a grave and considerable risk when taking action. If we can imagine, and I certainly believe we can, that in a world of immortality, there will be more novel forms of torture. I can imagine taking any action against oppressive or harmful entities whilst knowing that one risks excruciatingly painful torment that this is also a great act of courage. I would even be willing to put forth the idea that in such cases an immortal might be even *more* courageous than a human, as while a mortal knows that they will inevitably succumb to death, any excruciating pain they encounter will only be temporary. This might seem to be offer meager comfort to someone experiencing extreme suffering, but suffice it to say it will at least provide them with some eventual relief from their plight. But an immortal, chained to a rock and subjected to the painful act of having his entrails repeatedly eaten by an eagle, over and over again, has no recourse to death as an eventual end to his pain; he can be left in such a state for all of eternity.

Nussbaum also states that a lack of death would result in a rupture of the foundations on which we judge all of our relationships and values, such as friendship, as a willingness to die for the beloved is part of what constitutes that love. I for one do not find it particularly salient as it relates to friendship to measure my friend's love for me by whether or not they are willing to risk death on my behalf; to me, values like love, affection, and care for one's peers are not merely limited to vulnerability prescient in death, but relate more significantly to a desire to show respect,

consideration, and interest in the lives of people around us because we care deeply about them. To expect implicitly that they sacrifice their lives solely for me would be incredibly selfish on my part and would fail to consider all those who might love and need *them* besides *me* alone. It is true, however, that the willingness to make a great sacrifice would still be a major facet of love and affection, but these are *byproducts* of our love for one another, not necessary conditions of it, and it is not immediately clear than an immortal life removes any of these considerations nor the possibility of considerable sacrifice at all. We are still dealing with *humans* who have become immortal here, and I doubt the psychology of love in such a case would differ drastically from what we encounter today. One can still put oneself in harm's way for a loved one even as an immortal; the fact that there is no risk of a loss of life does not seem to reduce the sort of powerful love evident in such an act.

Similarly, many proclaim that a loss of death would mean a loss of urgency, as our approaching death typically puts a fire underneath us to achieve goals which we would otherwise put off. This value is regarded as compromised when death is removed from the picture. But this need not be so. Take the following example as an illustration of why urgency might still be a part of an immortal human experience: An immortal, let's call her Eleanor, has recently learned that another immortal for whom she's carried a longstanding flame, let's say, Diane, has recently been widowed after a 500-year long marriage due to her partner choosing to die and is now seeking a new life mate. Diane is exceptionally attractive and pleasant as a person to be around; however, she tends to confine herself at home with her partner, and as such Eleanor's access to her is quite limited. Diane is a serial monogamist and she has decided that once she marries her new partner, that even should they choose to die she will never remarry again. Take it for granted that Diane is in many ways quite the stubborn character and that her commitments have stood strong and unflinching in the eons she has spent on Earth; after all, the only thing that ended her previous marriage of 500 years was her partner's sudden decision to pass away. Eleanor knows it is highly likely that Diane will soon find a new suitor that fits her interests and can replace the last, as she is a desirable mate and has many prospective suitors; she will understand the sense of urgency required of her for her to take advantage of the opportunity of her beloved now being available. Take it for granted then that Eleanor's tastes are specific enough that there will be none she loves like Diane ever again, as evidenced by the fact that for the past 2,000 years, her interest in Diane has not budged in the least. If Eleanor is wise, she will look at the situation with urgency, as she

cannot let the opportunity pass in hopes another love will arise, as in the past thousand years no one has compared to her relationship with Diane. As such, there do seem to be feasible scenarios in which urgency is still a fundamental part of human experience.

Granted, there are some activities which would seem to not be urgent in such a life, and some activities would indeed be pushed off in favor of doing them another day; but it's unclear how this presents that much of a problem for an individual considering immortality. Urgency seems to be necessitated by time limitations but urgency in and of itself seems to present no real benefit to the urged other than obtaining a certain experience more quickly than one might have initially. And to an immortal, it seems there is no real reason to lament having not performed an activity sooner rather than later, provided they know they can still accomplish it in the future. It seems then that we ought to be skeptical about whether or not a loss of urgency is truly a loss at all in some respects. And as I posited before, it is easy to imagine that there might be other activities which crucially still have time limits on them. Say one wants to catch a particular type of fish which is very difficult to find. They certainly do not have a thousand years, let alone even a million, because evolution, environmental changes, and various other considerations will still need to be taken into account in how long they wait to do so. They might miss out on the fish if they continue putting off every fishing activity for an eternity. What we can however say is that immortals might prioritize doing things much more slowly than mortals do (though I can imagine things like deadlines will still be imposed on some activities), and this hardly seems to problematize a desire for immortality provided one is sufficiently patient. Provided they are not, well, then they themselves are responsible for ensuring they operate in a timely manner; I am also not convinced that the prescience of death stands to urge a procrastinator to accomplish things any sooner even in a mortal lifetime.

Now, there is one thing I will grant Nussbaum. It is true that the immortals might develop a different code of ethics and values based on their experiences; but in as much as they are still human, and they have not become completely alienated entirely from our understanding and experience of the world, wanting to be immortal is a rational wish, and being immortal itself does not seem to present a *severe* detraction from one's Earthly values. However, as we will see later, I will also argue that becoming totally alienated from our experience as humans should not at all be a source of anxiety provided certain conditions have been met, and that this should not concern us as much as philosophers like Nussbaum may posit.

In the following section, we will continue on in our consideration of urgency, as well as considering how things like innovation, creativity, and experience will be affected by an immortal existence, splitting it into two camps: Fixed abilities, in which our limitations are still present in that there is a threshold past which we cannot reach in achievement; and non-fixed abilities, in which this threshold is destroyed and humans overcome their Earthly boundaries.

SECTION III: INNOVATION, FIXED ABILITIES, AND NON-FIXED ABILITIES

3.1: Fixed Abilities

Philosophers such as Aaron Smuts (2011) have argued that an immortal life of fixed abilities, or a life in which our bodily and psychological limitations are still present, would present us with a challenge as it relates to innovation and long-term satisfaction. Smuts argues that it is feasible that in an immortal life, there will be some form of recurrence in terms of historical events. If that's the case, then any opportunity would not be held with the same weight, as we would have no urgency about an opportunity that would arise again anyway in the future. I find this line of argument deficient primarily because it fails to consider both our phenomenological experience of time as well as pain, as I will show in the next paragraph, and because, as I have shown in previous sections, it is still likely that some historical events will be truly unique and impossible to replicate or experience again given certain time constraints. This view presumes that somehow the immortals have drastically different conceptions and experiences of time, and considering they will still be functioning in human forms with human brains, I am not sure if I find this line of argumentation convincing. I doubt the phenomenology of their experience would be that much drastically different here.

As a direct byproduct of this, there is no reason to suppose that the immortal wouldn't seek to limit their experiences of pain to short durations merely because they can always do it later. There are still risks for an immortal being, as such, because one is still experiencing the regular passage of time; even if 1000 years seems short to an immortal, we can presuppose that experiencing the passage of time in such a time period would still be extremely significant and weighty. 1000 years is no small chunk of experience as a human. Even if life seems to move faster to us as adults, we still all experience the same year in duration when it comes down to it, and a year filled with joy is still preferable to a year filled with pain or despair, no matter how long one

is set to live. Or, to bring Eleanor back, Eleanor might be dreadfully romantically lonely and desperately desire a romantic partner. To her sensate system, it is perfectly reasonable to desire a lover *sooner* as opposed to *later*, which would lend her decision to find a lover extreme significance. The fact that, in another thousand years, she *might* be able to find a lover whom she is compatible with again other than Diane is little comfort for someone who is still in the midst of experiencing extreme loneliness for such a long period of time.

Continuing on, Smuts argues that even if we ought not to take it for granted that there would be an eternal recurrence of the same events, and as such there might still be a prescient urgency available to immortals to exploit, inevitably we will reach a point in which there is just nothing left to be done. He says: "To see why, imagine an infinite library or an infinite gallery after an infinite amount of time, even if you could find space on the shelves for a new volume, even if you could find readers for your millionth novel, why would you care to write it?" (Smuts 2011, p. 144).

I think Smuts's main problem has to do with a lack of consideration regarding our various reasons for attempting a project. Focusing on innovation as a means of providing meaning and value to our lives would obviously fall flat if we ever reached a point in which there is nothing else to innovate. But many of us find meaning and value in our lives beyond our novel contributions to intellectual or creative pursuits. Curiously, it becomes evident through a reading of his paper that Smuts takes for granted a prioritization of novelty in projects. He says: "Indeed we may never reach a point where all the love songs have been sung, but we can reach a point where there are no significant new kinds of love songs to sing" (ibid., p. 144). But art is not made solely for the sake of creating something *new* in the world. The creation of a new love song does not need to justify itself by demonstrating that it is a new type; if that were true then it seems that as humans all we truly prioritize is novelty, and there seems to be more at stake here. Sometimes it is about expression, creating deep meaningful complex stories, sometimes it is about wanting to experience something magical and deeply sustaining to one's existence, and sometimes it is about doing it merely for the sake of doing it. A good love song is not just written with the intent to create something new, but to personally express one's affection for another in a manner that moves their lover emotionally or others who hear it. Who cares if all the love songs and poems that have ever been written have been written? If I had a talented immortal musician as a lover and he deigned to merely reach into the infinite pile of poems and songs available and fling one in my

general direction I would consider it lazy indeed, but if he wrote one such song, specifically for me, with my features and drawing forth from his affection, I would be overjoyed! It would not matter the slightest bit that perhaps there is an identical song he sings somewhere in the infinite archives of human knowledge; what would matter to me deeply was that it came straight from *him,* that he made the effort, that he was the one who created it.

Now, there is one particular thing I will grant Smuts, and that primarily has to do with the fact that all things considered, an immortal life of fixed abilities, in which an individual will eventually reach a threshold past which they can no longer develop, say morally, creatively, athletically, etc., that such a life seems to carry with it great risks of eventually becoming frustrated by their limitations. From this perspective, I would like to argue that an immortal life of non-fixed abilities, or one in which due to technological developments or some other mechanism we manage to surpass our human limitations, would be very desirable in comparison. This is not to say an immortal life of fixed abilities is itself not desirable in comparison to the alternative, namely a death in which everything one is, including one's consciousness, personality, and functioning body, cease to exist and there is no longer any more experiential possibility, but that an immortal life of non-fixed abilities is largely preferable to a life of fixed abilities. With immortality at our immediate disposal, the likelihood of our developing technological mechanisms to surpass these limitations seems inevitable.

3.2: Non-fixed Abilities

Smuts, however, still argues that an immortal life of non-fixed abilities would render us into something like omnipotent Gods, in which our limitations have been surpassed and our ability to achieve success is now practically guaranteed for all intents and purposes (ibid., p. 146). Still, he says, our development at such a point would be strained by the potentially conflicting interests of the other immortal beings around us, many of which will butt heads with our plans and present obstacles to our development or growth.

Firstly, human life as it is already necessarily involves human conflict, and yet we do not see individuals proclaiming a legitimate wish for death merely because they have butt heads with their peers on occasion. Secondly, assuming that humans at this point of development would not have developed some manner of conflict management skills that outpace our own underestimates the possibilities of advancement here. It seems dubious to state that this is an unrealistic

presumption of what is possible given our time constraints. Thirdly, and following from the first point, we have to consider that our embodied environment might be drastically different from the ones we encounter today. Take the following possibility regarding virtual reality suggested by Alphin Jr.: "… it seems plausible that we will ultimately be able to upload our minds into a simulated environment, or an experience machine, and then live out fantasies or create challenges in a manner that ignites categorical desires" (Alphin Jr. 2010).

The suggestion here is that immortal individuals could create a virtual environment in which virtually anything could be achieved; disagreements do not stand for much when you can essentially collaborate with only those who are mutually interested in the arrangement. If I create a world with a completely unique set of natural laws, any of my mutually interested peers would have the chance to join me should they take an interest in my projects; those who do not will merely not upload themselves to this particular virtual reality cloud, choosing to go their own way and develop along their own lines.

Smuts admits to the fact that if desires were in harmony, non-fixed immortal beings would truthfully be able to create and do anything; however, curiously, he says that such a state is still not desirable: "Although feeling one's powers expand can be a source of great satisfaction, one can only move a mountain or destroy a galaxy so many times before it loses its novelty" (Smuts 2011, p. 147). My immediate reaction to this is that the mere expansion of powers for the sake of destruction and intrigue seems to be an impoverished view of what is possible and why we might embark on such a journey to begin with. Besides the infinite possibilities available, I also think that we pursue things as experiential creatures. Our lives are not merely engaged in for the sake of obtaining certain end goals alone; we do not live *merely* for the sake of completing projects or achieving things. I for one owe a great deal of my desire to live to the possibility of *experiences*. To experience new things, to encounter new worlds, to experience the wonder and awe at my existence. An immortal life of nonfixed abilities could offer some fairly incredible experiences at our immediate disposal, many of which we might find dizzying now but which might be exhilarating if we ever had the chance!

We need to think a little more creatively here about the opportunities. Imagine being able to completely create a universe almost entirely from scratch; to create various models and lovingly watch your creations develop consciousness, to experiment and develop different modes of intelligence on various planets, to quite literally have the infinite nature of possibility at your

fingertips (or whatever form you end up taking). It excites me greatly to envision a possible world in which I could learn how to split my consciousness up and embody several different lifetimes at once, learning from each individually while at their end recombining back into a whole self that can then use the fruits of each existence to develop something completely new and beautiful. That would be quite incredible! I might not be able to comprehend such an entity from my current viewpoint, but I can imagine the experiences I would have would be quite profound, and valuable in their own right.

There is one respect in which I will grant Williams's and Nussbaum's critiques of immortality in this regard, and that is that in such a state we truthfully might not recognize ourselves as humans any longer. Such a state might be completely beyond our conceptual reach, our experiential nature having surpassed anything we can envision on Earth in our current corporeal forms; but this should invite curiosity, a wait-and-see approach as opposed to trapping ourselves in the limitations of our current constraints out of fear we might completely lose ourselves and our uniquely human values. It is still rational to say we can desire such a state despite the fact we could never truly understand what such an existence would entail, though we can reasonably be certain there would be great beauty and wonder at our disposal to experiment with and develop into various modes of existence. And should we inevitably reach a point of contention in which we have determined in our absolute boredom that nonexistence is preferable, we could always just pull the plug on ourselves and cease to be.

Now, it is also true that such a state would require great responsibility. None of us would want a particularly sadistic and murderous personality to gain immortality of the kind I'm suggesting here; it would be terrifying to envision what tortures such a being could unleash onto conscious beings purely for their own sadistic satisfaction. But provided there is a general code of ethics about whom this immortality is allowed, and provided that the immortal beings provide constraints to the development of their capacities (for example, they could provide checkpoints at every turn to ensure none of their peers are engaging in atrocious acts, we might be able to grant that their sadistic desires would inevitably be slighted by the forward-thinking of other immortals), I do not think we would need to spend excess time worrying about the potential pains such individuals would cause.

CONCLUSION

I hope to have demonstrated that a corporeal immortal life would be desirable, despite claims to the contrary. I have shown that identity arguments that attempt to constrain it reduce and limit desirability unnecessarily. I have shown that arguments regarding boredom seem unsubstantiated empirically as well as logically when the potentials for satisfaction and fulfillment are truthfully considered. I have argued that an immortal life need not result in the absolute loss of human values in a relevant way and that many human values would still be retained due to physical constraints and time constraints. Finally, I have argued that a life of fixed abilities is still desirable, but the most desirable life of all would be one in which our earthly limitations have been surpassed. We might never come to see immortality in our lifetimes, but should the blessed day come when humans no longer have to face our mortality as a rule, I hope this work will stand as an argument for its viability.

Bibliography

Alphin Jr., Henry C. (2010) "Singular Immortality: Desirableness Through Technology and Liberty," *SSRN Electronic Journal*

Altshuler, Roman (2015). "Immortality, Identity, and Desirability," in: Michael Cholbi (ed.), *Immortality and the Philosophy of Death* pp. 191-203 (Rowman & Littlefield)

Bortolotti, Lisa., & Nagasawa, Yujin. (2009) "Immortality without boredom," *Ratio* 22(3), pp. 261-277

Fischer, John M. (1994) "Why immortality is not so bad," *International Journal of Philosophical Studies* 2(2), pp. 257-270

Hesiod. (1983) "Theogony," in: A. Athanassakis (Trans.), *Hesiod: Theogony, Works and Days*, pp. 1-102 (The Johns Hopkins University Press) (Original work published c. 700 BCE)

Nussbaum, Martha C. (1989) "Mortal Immortals: Lucretius on Death and the Voice of Nature," *Philosophy and Phenomenological Research* 50(2), pp. 303-351

Overall, Christine. (2003) *Aging, Death, and Human Longevity: A Philosophical Inquiry*, pp. 124-182 (University of California Press)

Smuts, Aaron. (2011) "Immortality and Significance," *Philosophy and Literature* 35(1), pp. 134-149

Williams, Bernard. (1973) "The Makropulos Case: Reflections on the Tedium of Immortality," in: *Problems of the Self*, pp. 82-100 (Cambridge University Press)

THE PLANE OF IMMANENCE: DELEUZE AND OTHER HERETICS

Eric Phipps

INTRODUCTION

The following article explores the various intersecting lines of thought between Baruch Spinoza's *Ethics* and the work of Gilles Deleuze. Subsequently, an assessment of Alain Badiou's critical figuration of Spinoza's thought will be engaged in conversation with that put forth by Gilles Deleuze. Badiou's sustained effort to reduce Spinoza's ontology to a circular doctrine of causality, by positing an errant void at the level of the infinite intellect, fails to recognize the distinction between the equality of divine powers, the attribute of thought, the infinite intellect, and the finite modes. Spinoza's conception of immanent causality, as it pertains to substance, demands the institution of a univocal plane populated with his concepts of the attributes, the modes, the infinite, the finite, and divine nature. The presupposition of immanence at the center of the onto-ethical system in Spinoza's *Ethics* effectively prefigures the heretical form of philosophical deterritorialization and constructivism later adapted, modified, and elaborated by Deleuze in defining his plane of immanence.

SECTION I: SPINOZA'S HERETICAL IMMANENCE

As Spinoza shows us, there can only be one substance that comprehends all infinite and finite things. Substance is self-contained, self-conceived, and relies on nothing outside of itself for its formal existence. God is an absolutely infinite being consisting of infinite attributes. Each of the attributes expresses the eternal and infinite essence of God.[1] The modes, then, are the affections of substance, neither self-contained nor self-conceived. The essence of substance necessarily involves existence and it is necessarily infinite, which is to say that the nature of substance is such that it envelops all reality absolutely.

The more reality something has the more attributes it has, which entails that an absolutely infinite entity must consist of infinite attributes, each expressing a definite essence. Expressions of an attribute express the eternal, singular, and infinite essence of God. An absolute entity, such as God, must have absolute power to exist, and the degree of power attributed to such an entity finds its intrinsic determination in the degree of reality it has. Thus, the nature of substance produces in itself the absolutely infinite power and energy to exist, which admits of no external

cause for the energetic power to exist that substance possesses absolutely. To put it differently, no substance other than God can be conceived, and crucially, both thinking and extended things are either attributes of God or affections thereof. Finally, neither an attribute of substance nor an absolutely infinite substance itself can be divided. Spinoza maintains that it would be absurd to understand 'absolutely infinite' as determinable by a certain quantity of divisible parts. In other words, there can be no portion of infinity that can be numerically distinct from the whole of infinity. Absolutely infinite substance can only be conceived of as distinct in terms of its reality, that is, only in its *absolutely infinite reality* (Spinoza 2006, pp. 3-11).

To unpack some of the moves that Spinoza makes in the first half of the *Ethics* it is important to draw attention to a certain stylistic tendency in the presentation of his work. Many early definitional, axiomatic, and propositional reference points make appearances later in the book as constructive recapitulations for new purposes of conceptual development. A good reader of Spinoza ought to be looking for conceptual clarity in later proofs, corollaries, and scholia, rather than favoring a close reading of only a few propositions. On one hand, a reader must entertain a general overview of the work, yet cannot grasp the entire system at once. On the other hand, one must make careful selections of certain angles, curves, and folds within Spinoza's thought to help motivate the passage beyond traditional dialectical suppositions. Our hope for the current effort is to draw out the author's own interest in producing greater clarity and comprehension of the natural world by thrusting it into conversation with modern Spinozist texts. Spinoza appears to sustain a thorough interest in the bodies, lines, and planes of Euclidean geometrical theorems, and through this interest has produced many interwoven layers of intensional consistency within his systematic usage of definitions, axioms, propositions, proofs, corollaries, and scholia. I have done my best to provide a cursory account of the major components of Spinoza's ontological system. However, to understand the correlative linkages of this complicated system in its entelechy requires sustained commitment to the exegetical labor that Spinoza's text demands. In the interest of tracing some of the implications of Spinoza's ontology we will now examine his notions of finite being, causality, thought, and extension.

Firstly, Spinoza declares in E1Def2 that entities are conceived as finite beings if they are limited by other entities of the same nature (ibid., p. 3). To put it differently, finite being is defined by its externally shared surfaces, constitutive milieus, and environments of interaction with other things. External determinations are what establish the breadth of activity wherein a finite thing

partakes. Thus, every finite being is in some way or another caused to act in a certain manner through existing with others and being acted on by them. Insofar as God is the ultimate causal power in Spinoza's ontology, it follows that the finite modes are themselves determined and defined through their expressive relationship with the absolute divine nature that conceives both their essence and their existence (ibid., pp. 13-18). Therefore, finite beings refer beyond themselves for determination, and their individuation must occur through a process of adopting powers, persevering with others, and enduring external causality to becoming higher order individuals.[2]

> In reality, when the mode comes to exist through the action of an efficient cause, it is no longer simply comprehended in the attribute, but continues to exist…or rather tends to do so; that is, it tends to persevere in existing (Deleuze 1988, p. 62).

Finitude has duration, which entails that there is a beginning to finite being but not necessarily an end. Deleuze signals a radically different interpretation of finite being, yet his fidelity to the details of Spinoza's system remains remarkably consistent, even while constructing a conceptual system of his own in the process.

The dualism of infinity-finitude is typically conceived analogically as a continuous-discontinuous binary sense of quantity and duration in the world. Viewing this dualism through the perspective of finitude, as human beings, we may find that we cannot yet comprehend the infinite. In typical interpretations of mathematics, infinity is thought of as emanating from a process of continual accumulation of finite things until properties that no longer pertain to finite being have emerged. Alternatively, finite things may be thought to either approximate or limit the infinite in quantity. There may be a third position, however, which takes the infinite and finite as two distinct manners of being with their own structural frameworks. On this view, the infinite is conceived as enveloping the finite, while the finite affirms an expressive modality within the infinite (Duffy 2004, pp. 55-56). Mathematician Albert Lautman develops this notion of the infinite-finite relation, which places the demand of analyzing both domains of being from the perspective of their existing structural frameworks and systems of reciprocity (Lautman 2011, pp. 81-83). This expressive mapping of the infinite-finite relation will later be adopted by Deleuze to

develop a concept of Spinoza's modal essence in contrast to modal existence. Modal essence will be explicated in the latter portion of this article.

Secondly, E1P28 of Spinoza's *Ethics* establishes the relation between finite modes that results in their individuation through extrinsic inter-determination causing finite entities to exist and act in certain ways. This inter-determination should not be understood to describe a set of interactions between presupposed individuals, but rather an active process of causal individuation. Determination (*determinatio*)[3], for Spinoza, is closely tied to a sense of externally bounded corporeal limitation as a causal principle among the finite modes. Proposition 28 makes use of a familiar definition in further developing Spinoza's concepts of both causation of finite modes and their individuation (E1Def2).

> A thing is said to be finite in its own kind [*in suo genere finita*] when it can be limited by another thing of the same nature. For example, a body is said to be finite because we can always conceive of another body greater than it. So, too, a thought is limited by another thought. But body is not limited by thought, nor thought by body (Spinoza 2006, p. 3).

In considering Spinoza's conception of motion-and-rest in relation between finite bodies, Shein argues in favor of reciprocally active inter-determination between bodies in Spinoza's metaphysics, against a traditional notion of efficient causation. Finitude and limitation are associated with the Latin verb *terminare*, which signifies a primary notion of delineated boundaries and secondarily the relation of limitation, definition, or delimitation a thing (Shein 2015, pp. 335-337). In other words, a body placing a limit or boundary on another is the causal determinant in relations of modal existence. There can be no vacuum or void in Spinoza's metaphysics, therefore we cannot posit any unique surface retained to a certain finite entity, but rather finitude is defined by the manner of limitation enacted upon it within the shared surfaces enveloping finite bodies (Shein 2015, p. 350). Thus, there is a continuous nature to the network of inter-determinable bodies, i.e., nothing finite can be intrinsically determined, and it follows that every finite entity modifies an infinite attribute of God.

Thirdly, E2P7 supports equality at the level of the attributes of thought and extension, which is to say that the order of connection between ideas and things are the same insofar as their singular essences consist in the attributes of either thought or extension, both of which are

comprehended by one and the same substance (Spinoza 2006, p. 32). This discovery suggests that our preceding discussions of finite being and causality remain consistent across the known attributes with which human experience consists, i.e., thought and extension. What remains central to the equality of the attributes is that ideas and bodies, as finite modes, perform completely autonomous acts—both ontologically and epistemically. However, they maintain correspondence with one another through the expression of their attributes within the order and absolute power of substance (Deleuze 1988, p. 87). Thus, the doctrine of parallelism is held together by the absolute causal necessity of God's essence. Spinoza asserts that in order to understand this union of the mind and body we must develop an adequate understanding of the nature of the body, which constitutes the idea of the mind through *affections*[4] of the body (Spinoza 2006, p. 36). The mind simply is the idea of our body for Spinoza; not in a purely representational sense, but rather, "the *idea* that we are is to *thought* and to *other ideas* what the *body* that we are is to *extension* and to *other bodies*" (Deleuze 1988, p. 86. *Italics added for emphasize*).

Finally, Spinoza admits of no analogical conception between the order of creation and the order of things. Being is not conceived of as subject to contingency or possibility, but rather must involve the unified causal necessity of absolute divine nature. In fact, only the *univocity*[5] of the whole can be affirmed through *modal expression*[6] of substance, which *univocally* comprehends all that is in absolute nature. This essential univocity of being is supported by the scholium following E1P29, which declares that speaking of the overall system of nature refers to *Natura Naturans* on one hand, while on the other hand, speaking of anything that follows the necessity of God's nature refers to *Natura Naturata* (Spinoza 2006, p. 19) In other words, *Natura Naturans* involves substance and causality, while *Natura Naturata* involves the effect and modality of a conception of nature.[7] Deleuze asserts that these two orders of nature "are interconnected through a mutual immanence: on one hand, the cause remains in itself in order to produce; on the other hand, the effect or product remains in the cause" (Deleuze 1988, p. 92). Moreover, Spinoza's construction of a univocal plane of being serves to *deterritorialize* the equivocal stances of neutrality and ambivalence oriented toward the order of creation, and thus is able to liberate the concept of immanent cause from the constraints of presupposing a transitive theory of divine creationism (Deleuze 1990, p. 67).

In *Difference and Repetition*, Deleuze claims that univocity has always been the one-and-only ontological presupposition. "There has only ever been one ontology, that of Duns Scotus,

which gave being a single voice" (Deleuze 1994b, p. 35). The Parmenidean and the Heideggerian enunciation of being deploy the same fundamental univocity, according to Deleuze. Univocal being is the same across all modalities, yet the essence of univocal being also includes all modalities of difference within it. Thus, difference itself finds its determinacy in univocity, yet modalities of difference are not the same in essence, but are the same in respect to expressive univocity of being (ibid., p. 36).

Deleuze is interested in how Spinoza institutes his immanent causality at the center of a plane of nature. He declares that Spinoza was abundantly aware that thinking immanence, as immanent to something transcendent, would ultimately constitute an inadequate understanding of nature. Deleuze is mostly interested in how Spinoza conceives of the finite modes as constituted by an immanence of being. He expresses the interest in turning substance upside down, so that it turns on the order of the modes. It is in this creative process of inverting Spinoza's substance, and conceiving of an elaboration of the modal essences, that Deleuze institutes a plane of immanence in which finite modes are capable of expressing immanence at the center of being (Deleuze 1990, p. 11).

> There he attains incredible speeds, with such lightning compressions that one can only speak of music, of tornadoes, of wind and strings. He discovered that freedom exists only within immanence. He fulfilled philosophy because he satisfied its prephilosophical presupposition. Immanence does not refer back to the Spinozist substance and modes but, on the contrary, the Spinozist concepts of substance and modes refer back to the plane of immanence as their presupposition (Deleuze 1994a, p. 48).

SECTION II: THE PLANE OF IMMANENCE

In order to better situate our discussion of Deleuze's plane of immanence we will briefly outline some key incommensurabilities pertaining to the adjacent practices of philosophy and science. Certain *zones of indiscernibility* exist between these practices, due to the variable approaches that they present in considering chaos, infinity, and first principles. Thus, the diagrammatic features or topographical mappings of particular fields of interest may differ. We will sketch what Deleuze calls the "three heads of opposition" between science and philosophy below:

(1.a) Philosophy institutes a *plane of immanence*

(1.b) Science provides a *system of reference*

(2.a) Philosophy deploys *inseparable variations* and *concepts*

(2.b) Science employs *independent variables* and *functions*

(3.a) Philosophy expresses the event through *conceptual personae*

(3.b) Science determines states of affairs through *partial observers*

Firstly, how shall we distinguish *partial observers* from *conceptual personae*? Deleuze designates these terms as *sensibilia* respective to each practice. *Conceptual personae* are understood as the *perceptions* and *affections* of disparate concepts, while *partial observers* are said to be the *perceptions* and *affections* of scientific functions. *Sensibilia* in philosophy are what make concepts capable of being perceived with a certain feeling or mood at infinite speed. *Sensibilia* in science, on the other hand, seem to allow functions to perceive systems of reference at slower speeds, thus feeling a sort of finite biomechanical vitality of independent variables (Deleuze 1994a, pp. 129-133). Thus, conceptual personae can be grasped as the various philosophical *masks* donned by a particular thinker. The dramatic presentation of concepts on an infinitely variable plane enacts the intensive character of a particular thought, increasing the affective force with which a concept is felt. Partial observers, on the other hand, are sensory affections and points of view in the functives themselves. The perceptive notion of partial observers relates to the quantity of information employed within a system of reference, while its affective notion is conceived as an energetic relation between elements of scientific propositions (ibid., pp. 129-133).

Secondly, we must distinguish scientific functions from philosophical concepts. A concept, for Deleuze, must be created upon the *deterritorialized* grounds of constantly new planes of *consistency*. In explicating his notion of *geophilosophy*, Deleuze draws our attention toward a revolutionary motivation intrinsic to concepts and their events. He asserts that the *event* of joyful '*enthusiasm',* standing beside the *state of affairs*, enables an "absolute deterritorialization" of historical limitations to the extent that a people can call for a "new earth". Invoking Kant, Deleuze declares that a revolutionary *utopia* always stands for an *absolute deterritorialization*, placed at a critical point of association with a present relative milieu. Thus, *utopia* designates, in the *event*, the constitutive conjunction of the philosophical concept with the specifically repressed forces of

an exigent psychosocial or political milieu. The self-positing nature of revolutionary enthusiasm is what enables its apprehension in past, present, and future forms.

> …*enthusiasm* with which men embrace the cause of goodness…is always directed exclusively towards the *ideal*, particularly towards that which is purely moral…even the old military aristocracy's concept of honour (which is analogous to enthusiasm) vanished before the arms of those who had fixed their gaze on the *rights* of the people to which they belonged, and who regarded themselves as its protectors. And then the external public of onlookers sympathised with their exaltation, without the slightest intention of actively participating in their affairs (Kant 1991, p. 183).

Hence, the philosophical concept does not depend on an objective reference point within the present *state of affairs* to be apprehended. Rather, it is a formal (or territorial) distinction, marked by the *event* of immanent enthusiasm, which frees immanence from all external limits imposed upon it in the objective *state of affairs*, i.e., by modern political economy or capital itself (Deleuze 1994a, pp. 99-101).

Scientific functions, on the other hand, trace the independent variables of a given *state of affairs* from the outside. The goal of science is to creatively extrapolate *prospects*[8] onto a system of reference (Deleuze 1994a, p. 24). Functions constitute scientific, logical, and doxological propositions (ibid., p. 155). The history of philosophy is not the succession of systems, one after the other. Rather, Deleuze characterizes the history of philosophy as a becoming and coexistence of interleaved planes of immanence (ibid., pp. 57-59). Reconstituting philosophies of the past is always a process of selection, mediation, institution, and creation. Assuming the image of an earlier thinker must involve the extension of the original plane, while coextensively establishing new coordinates between older and newer planes. Thus, the demonstrable irreducibility of concepts to discursive systems, expert opinions, and unwavering veridicality means that a concept must be distinct from a function, insofar as particular concepts are distinguished by intensive ordinates on their plane of immanence and functions by their extensive coordination with their system of reference (ibid., pp. 57-59).

Lastly, our remaining task is to distinguish a plane of immanence from a system of reference. The plane of immanence itself is a regional sectioning of chaos that "acts like a sieve"

(Deleuze 1994a, p. 42). It is through the correspondences of atmospheric chaos, corporeal bodies, and thought appearing in the formation of porous metastructures that a plane becomes capable of retaining consistency. Over multiple processes of interleaving planes, each with its unique conceptual stratification and elemental diagrammatic features, the problem of philosophy is to acquire consistency without losing the infinite into which thought plunges (ibid., p. 42). Deleuze directs our attention back toward the first philosophers, thus reminding us that by stretching immanence "like a sieve" over chaos, the Greeks were unique in placing an immanent order to a cosmic milieu. In the Greek moment, the first philosophers institute a plane, upon which infinite movements pass into philosophical consistency on two sides: *physis* and *nous* (ibid., p. 44).[9]

The *image* of thought finally arrives on the same plane through *nous*, while *physis* endows being with *substance*. In pre-Socratic thought, the diversion of wisdom toward the same plane constitutes a unique moment in the history of philosophy. This constitutive diversion has the power to put philosophy at the service of a *pure immanence*.[10] Summoning the infinite to the center of being and thought, the pre-Socratics provide a fertile topography for the creative act of philosophy. Deleuze asserts that we witness the most rigorously drawn distinctions between *physis* and *nous* in Anaximander, yet always maintained on the same plane (ibid., pp. 43-44).

> Abstractions explain nothing, they themselves have to be explained: there are no such things as universals, there's nothing transcendent, no Unity, subject (or object), Reason; there are only processes, sometimes unifying, subjectifying, rationalizing, but just processes all the same (Deleuze 1995, p. 145).

The tendency to conceive of the human intellect, at every turn, as capable of peeling away from nature only seems to diminish our understanding the processes that we must endure to live our lives purposefully. Imagination and actuality share an interesting correspondence with one another in Deleuze's thought, as well as Spinoza's. The image of thought seems closely related to the affections of the body for these two thinkers, which remains a compelling way to view thought. Spinoza's sense of adequate and inadequate ideas plays a crucial role here. The more we understand causal determinations that limit our bodies and their affections, the more we tend to produce adequate ideas, and it seems that this active orientation toward the body ought to be capable of drastically altering social conditions and bonds of friendship. Remaining reflective of

the pre-Socratic moment in philosophy, much of the force in Spinoza's work is in the way that the author continues pushing the work towards its fullest and most robust logical expression of the plane of nature.

It is Spinoza who finally reactivates this pre-Socratic moment with his demonstration of a monistic plane of consistency—retaining the infinite speed of thought and being—without relegating anything to the outside. In Spinoza, thought is not conceived as a manner of abstracting oneself from the plane of nature. Rather thought as becomes a power of striving through the affects of the body to better form adequate ideas of nature, while combining with others to express absolute nature in its infinity (Beistegui 2012, p. 105). There always remains something of a heretical undercurrent in affirming the absolutely infinite as immanent to itself. Immanence is dangerous, although it is not immediately apparent why. As Deleuze tells us, immanence "engulfs sages and gods. What singles out the philosopher is the part played by immanence or fire" (Deleuze et al. 1994a, p. 45). A plane of immanence does not admit of transitive causes, which poses the most heretical alternative to vertical orders of creation, deduction, and our general desire to trust in mythical subversions of our intellect, and herein lies the danger for sages, gods, and priests. There are no reasonable, rational, or respectable methods for laying out a plane of immanence. Rather the lack of method relative to this effort relegates it to esoteric experiences and pathological processes, as Deleuze proclaims:

> These measures belong to the order of dreams, of pathological processes, esoteric experiences, drunkenness, and excess. We head for the horizon, on the plane of immanence, and we return with bloodshot eyes, yet they are the eyes of the mind. Even Descartes had his dream. To think is always to follow the witch's flight (ibid., p. 41).

SECTION III: THE HALTING OF THE INFINITE

Spinoza's conception of infinity has remained extraordinarily relevant within the fields of mathematics and physics. Looking no further than Leibniz or Malebranche, the physico-mathematical correspondence to Spinoza's geometrical presuppositions is obvious. Likewise, Deleuze's thought shares critical correspondence with geometry, calculus, and set theory that moves beyond the typical coordinates engaged in such works.

Turning to a discussion of Badiou's mathematical ontology and its interrelations with philosophies of immanence at this juncture will help to demonstrate the notion of a plane of immanence through a contrasting viewpoint of infinity.[11] Alain Badiou offers a treatment of mathematical sets in support of an immanent ontology opposing that of Spinoza and Deleuze, while imposing his own notion of the void, the subject, and the event as necessary conditions of an adequately immanent ontology. Badiou's critical project characterizes Spinoza's foundational work in the *Ethics* as a *closed ontology*, which is consistent with his repeated claims that both Deleuze and Spinoza fail to allow space for the void, posit no theory of the subject, and offer no framework for sufficiently conceptualizing events.

Firstly, Badiou claims that Spinoza posits an ontological circularity at the level of both the infinite and the finite. More precisely, the singular operations of the intellect mediate a complex schema of attributes and modes within the metastructure of the supreme 'there is', or the indexical term preferred by Badiou to refer to Spinoza's God. On this view, the infinite intellect is understood to produce a relation of excess that goes beyond Spinoza's strictly immanent causality. Citing E2P21, Badiou claims that this relation of excess results in a coupling operation that short-circuits identity under the attributes.[12]

> Like the relation of causality, the relation of coupling implies the existence of an infinite regress. Thus every mode has a cause, which itself has a cause, and so on. Similarly, every idea coupled to its object must be the object of an idea that is coupled to it (Badiou 2015, p. 93).

Despite Spinoza's ostensible preclusion of all excess from ontology Badiou discovers a coupling relation that may reproduce ideas of ideas and perhaps even produce replications of the ideatum, or the object of a particular idea, ad infinitum. In other words, Badiou implies that an emergent crisis of identity occurs at the level of the attributes through the infinite intellect. Such as in the case where an idea of the body must struggle to comprehend the disjunction between the attributes of thought and extension simultaneously. This would seem to illustrate the emergence of distinct coupling relations grasped by the intellect, which constitute distinct identities composed of oddly intellected mixtures of the "mind as idea of the body and the idea of the mind as the idea of the idea" (ibid., pp. 93-94).

In sum, Badiou locates the point of an infinite regress in Spinoza's concept of immanent causality, which emanates through God's substantial nature, and then he proceeds to interrogate the infinite intellect as some kind of quasi-substantial subject propagated through a coupling operation between ideas, bodies, and various attributes of God. Badiou describes this coupling operation of the intellect in the following manner:

> Ultimately, coupling 'infinitizes' every point of the intellect, just as causality 'infinitizes' every point of the 'there is'. … Undecidable in terms of its existence; atypical in terms of its operation; eliciting a doubling effect – these are the traits which, in my eyes, identify the intellect as a modality of the subject-effect (ibid., p. 99).

Secondly, the *undecidability* of the intellect is a term related to Badiou's insistence that the void must exist between the infinite and the finite, which is to say that there is a possibility that no causal realm of intelligibility lies between infinity and finitude. Another resulting possibility of the void is that our logical proofs for causal events are founded upon a region of non-existence, which would mark an inconsistency in the concept of infinite modes and call into question Spinoza's ontological claim that no vacuum exists in nature. On this view, the term *infinite mode*[13] is a nominal distinction, such that it remains deficient of its supposed referent (Badiou 2005, p. 119). To put it differently, we can neither empirically nor rationally verify the existence of infinite modes, as they are not given in experience and cannot be deduced a priori. On Badiou's analysis, Spinoza's infinite modes appear to prove his system conceptually—before disappearing ontologically into the "causal abyss"—swallowing its referent and leaving only an errant name in its place. Therefore, the infinite intellect, such that it is an infinite mode, depends upon the associated axiomatic system of proof put forward by Spinoza (ibid., p. 120). This is what Badiou means in declaring that the infinite intellect relies on the undecidability of the infinite modes in Spinoza's closed ontology (Badiou 2015, p. 94).

Thirdly, Badiou finds that Spinoza has left an untenable asymmetry at the level of attributes, which likewise entails an asymmetrical power relation of the infinite intellect—or the infinite mode of thought—to the other attributes. For Badiou, this undermines the supposition of equality at the level of the attributes that must exist a priori to any determinate mode. However,

the equality of both the attributes and the powers of God are demonstrated clearly by Spinoza at E2P7 and its corollary:

> *The order and connection of ideas is the same as the order and connection of things*…Hence it follows that God's power of thinking is on par with his power of acting. That is, whatever follows formally from the infinite nature of God, all this follows from the idea of God as an object of thought in God according to the same order and connection (Spinoza 2006, p. 32).

Badiou claims that the infinite intellect forces an exception to the above proposition insofar as it remains impossible to conceive of an inter-attributive mapping[14] from the perspective of the intellect itself. Badiou maintains that the attribute of thought cannot be isomorphic with any other attribute under this ontological arrangement (Badiou 2015, p. 94). The attribute of thought, according to Badiou, shares no isomorphic point of nexus with any other attribute, thus creating a sense that the intellect is foreclosed upon by Spinoza's own ontology (Craig 2013, p. 233).

Badiou's sustained effort to reduce Spinoza's ontology to a circular doctrine of causality by positing an errant void at the level of the infinite intellect fails to recognize the distinction between Spinoza's equality of powers, the attribute of thought, the infinite intellect, and the finite modes. In the next section we will respond to Badiou's critique by offering a defense of Spinoza's ontological presuppositions, while demonstrating Deleuze's fidelity to Spinoza's onto-ethical system with his adaptation, modification, and elaboration utilized to define his plane of immanence.

SECTION IV: A DEFENSE OF THE INFINITE MODES

Deleuze offers precision and clarity to an analysis of Spinoza's ontological system by distinguishing modal essence from the attributes and modal existence, while concomitantly explicating the role that individuation plays in finite being. If Deleuze's *modal expression*[15] argument stands, then it is clear that Badiou's account errs insofar as it equates the notion of modal existence to that of modal essence, and then offers a reductive conception of Spinoza's strictly causal relation of modes to the attributes. By omitting a deeper analysis of the role that the attributes play in the infinite intellect, and focusing much of the force of his argument toward

concerns regarding infinitesimal denumerability and the void in Spinoza, Badiou misses an opportunity to demonstrate his interpretation of Spinoza affirmatively (Roffe 2007, p. 404).

According to Deleuze, "a modal essence should be singular in itself, even if the corresponding mode does not exist" (Deleuze 1990, p. 196). Nonetheless, the path into this problem is not abundantly clear. In order to draw out the necessary distinction involved in making this connection between modal existence and modal essence, Deleuze turns to a brief comparison of Spinoza and the priest theologian of the High Middle Ages, Duns Scotus. In his notes, Deleuze tells us that the comparison only ventures insofar as to develop a thematic engagement with Scotus' notion of degrees or quantities of intensity, and not to a comparison to Scotus' individuation (ibid., p. 380). As Deleuze demonstrates, modal essences are distinguished from their attribute by their intrinsic degree of intensity, even in cases with no corresponding finite modal existence, while the attributes are distinguished by their quality. This comparison to Scotus is outlined as follows (ibid., p. 196):

(1) Whiteness has various intensities
(2) These intensities are not added to whiteness as one thing to another thing
(3) Adding intensity is unlike adding a shape on a wall by drawing upon it
(4) Degrees of intensity are intrinsic determinations and intrinsic modes
(5) Whiteness remains the same univocally under whichever modality it is considered
(6) Consequently, Spinoza's modal essences are intrinsic modes or intensive quantities

The important distinction beginning to take shape here is that which delineates both the inward and outward ontological folds determining the essence and existence of modes. Deleuze goes on to explicate this distinction by deepening the Scotus analogy before discussing Spinoza's unique concept of individuation.

Contrary to Badiou's claims, the attributes retain their univocal identity and admit of no asymmetry, while resisting any formal modification and containing all degrees of their affect. Although modal essences may be contained in an attribute they can be distinguished from them insofar as they are intensities of their attributive qualities, and each modal essence can be distinguished by a differing degree of intensity. Having a differing degree of intensity entails a modal essence attaining a specific intensive quantity at the level of being. Thus, "Only a

quantitative distinction of beings is consistent with the qualitative identity of the absolute" (ibid., p. 197). Consistency is key in this register, as the real and numerical distinctions share a certain constitutive correspondence between one another. Further explication by Deleuze shows that a quantitative distinction in modal existence actually constitutes an intrinsic difference of intensity. Moreover, every existent mode, or finite being, actually expresses the absolute according to this intensive quantity that constitutes its very essence (ibid., p. 197).

Proclaiming, as Deleuze indeed does, that finite being expresses the absolute is no mere figurative turn of speech, but is an actual concrete relation of expression of God's essence vis-à-vis modal essences within the inner fold of the attributes. First of all, the attributes must be understood as infinite and indivisible qualities that express the essence of substance (Roffe 2007, p. 402). Secondly, a modal essence, whether or not it exists as a finite thing in the world, must be understood as contained within an attribute. Finally, the intensive quantity that constitutes a modal essence is the degree of power with which it consists in the infinite attribute. Thus, a modal essence is intrinsically distinct both in quantity and intensity, which Deleuze relates to what he also finds in Spinoza's concept of individuation. "Intensive quantity is infinite, and the system of essences an actually infinite series" (Deleuze 1990, p. 197). Each modal essence is involved with the production of the infinite series of essences, yet each is an irreducible and singular unity within a concrete system—a "system of 'complication' of essences" (ibid., p. 198).

On the other hand, we discover modal existence in the work of Spinoza to be determined by extrinsic causal relations. Existence, and cessation of existence, of finite modes are determined by laws that remain external to their modal essence. As Spinoza shows us in E1P24 and its corollary, "whether things exist or do not exist, in reflecting on their essence we realize that this essence involves neither existence nor duration. So it is not their essence which can be the cause of either their existence or their duration, but only God, to whose nature alone existence pertains" (Spinoza 2006, p. 18). An existing mode must have duration and must also be extrinsically motivated in order to become individuated, which is to say that what determines existing things are other things of a similar kind within the environment that must be endured. Bodies, for example, are defined in existence through extrinsic mechanical laws and proportionality in regard to motion-and-rest, while minds are said to exist and act in a similar determinate fashion (ibid., p. 21). In other words, our thinking is also comprehended as finite, and is thus constrained through

extrinsic mechanical laws of God's nature, as evidenced in E1P30 and the second corollary to E1P32.[16]

To be absolutely clear, both Badiou and Deleuze are guilty of generously elaborating upon Spinoza's philosophy to make amenable their own figuration of immanence. Nonetheless, it appears that Deleuze retains a level of fidelity to Spinoza that cannot be located in Badiou's work. Deleuze seems to engage with the concepts as they are developed in Spinoza beyond the singular propositions, whereas Badiou misses developments as Spinoza modifies and explicates various elements of his immanent system in the *Ethics*. One aspect of Badiou's work that becomes immediately obvious is his emphasis on locating a mathematical ontology that accounts for infinity. Badiou would later reflect on this pursuit with the following statement:

> The dream or idea of a purely mathematical philosophy, which would be written exclusively with letters and symbols and in which all sentences would be theorems, is impossible. Spinoza represents the extreme case, having written his famous book *Ethics* in the form of Euclid's great treatise. But even Spinoza is very far from a rigorous symbolization (Badiou 2022, p. 40).

Thus, it appears that the figure of Spinoza signifies some kind of disappointment for Badiou, insofar as he attempted to construct an extremely robust geometrical ethics, yet perhaps failed to meet the demands of formal demonstration in theorems of mathematics. He then declares that philosophy relies on the meaning of words and not on the letter, which entails that it lies between mathematics and poetry (ibid., p. 40)

However, by putting so much effort into demonstrating Cantor's set theory in a purely mathematical ontology Badiou seems to have skipped over establishing a clear outline of Spinoza's concept of finitude. Rather, it appears that Badiou largely takes the notion of finitude for granted as he helps himself to a robust critique of both Spinoza and Deleuze. Finitude, for Spinoza, means something that is extrinsically limited, which makes no notice of denumerability of finitude. In fact, one could interpret Spinoza's finitude as infinitely composed since the world is constituted by a singular substantial infinity (Roffe 2007, pp. 398-400). Therefore, ontologically speaking, finite things are non-denumerable and infinitely interchangeable compositions defined by a relation to a constrained freedom to act and be acted on by other limits. We can conceive of finite

individuals limiting other finite individuals, yet there is no void that delimits the modal capacity to express the infinite as Badiou would have it. Infinity cannot be opposed, and it is precisely this that establishes the univocality of being and the unique nature of Spinoza's substance (ibid., pp. 398-400).

The attribute of thought, God's power to think, and the infinite intellect are all separate and distinct terms in the work of Spinoza, yet Badiou fails to recognize them as such. The infinite intellect is Spinoza's term for the infinite mode of thinking. An infinite mode is what Deleuze calls *modal essence*, which is contained within the attribute itself. Modal essences are conceived as the expressive essence that finite modes are capable of expressing.

> Yet *the distinction of powers and attributes has an essential importance in Spinozism*. God, that is the absolutely infinite, possesses two equal powers: the power of existing and acting, and the power of thinking and knowing (Deleuze 1990, p. 118).

Due to the conflation between the concepts of *infinite intellect* and *attribute of thought*, Badiou's critical project loses much of its force. The confusion between the *attribute of thought* and the *infinite mode of thinking*, i.e., the infinite intellect, is a fatal misreading. Understanding the distinction between God's power of thought (*Absoluta Cogitatio*), the attribute of thought, and the infinite intellect is crucial to a productive reading of Spinoza's *Ethics*. Badiou's critique effectively excludes the *Absoluta Cogitatio* by ignoring the priority of equal powers as outlined in Spinoza (Craig 2013, p. 238). Therefore, a more nuanced analysis at the level of God's power of thought, the attribute of thought, the infinite intellect, and finite modes in Spinoza's work have not been adequately dealt with in Badiou's critique.

CONCLUSION

Spinoza's conception of immanent causality, as it pertains to substance, demands the institution of a univocal plane populated with his concepts of the attributes, the modes, the infinite, the finite, and divine nature. The presupposition of immanence at the center of the onto-ethical system in Spinoza's *Ethics* effectively prefigures the heretical form of philosophical deterritorialization and constructivism later adapted, modified, and elaborated by Deleuze in defining his plane of immanence. Alain Badiou's critical figuration of Spinoza's thought and reduction of Spinoza's

ontology to a circular doctrine of causality fails to recognize the distinction between the equality of divine powers, the attribute of thought, the infinite intellect, and the finite modes. In this article, we have sought to demonstrate the failure of Badiou's 'closed ontology' critique of Spinoza on account of his fatal misreading of the *Ethics*. Positioning Spinoza as the anticipatory figure of Deleuze's notion of the plane of immanence enables a more adequate and stable account of what an elaboration of Spinoza's onto-ethical system of interleaved planes actually signals within the philosophical register, beyond Badiou's mere reductionist stance.

Notes

1. Spinoza's concepts of God, Substance, and Nature are closely intertwined. I am using these terms somewhat interchangeably, yet the context dependent appropriateness of each one of these terms has been followed to the best of my knowledge.
2. By higher order individual I mean a composition of finite modes with a higher degree of power due to its quantity of modal assemblages, which is conceived as having greater capability to act as a collective unity of bodies.
3. *Determinatio* as a limit, not as in determinism.
4. Affections (*affectio*) understood as the affects of the body, or sensory experiences of materiality, relating to the generation of ideas and the mind itself.
5. Univocity understood as expressing only one substance in terms of *what is*.
6. Modal expression understood as modalities of definite attributes that express God's essence.
7. *Natura Naturans* conceived as self-caused nature-in-itself, or substantial nature; *Natura Naturata* conceived as nature in perception, presented and represented by effects, related to the modes.
8. *Prospects* understood as Deleuze's term for functions, and variables. Designates the various elements of a scientific proposition.
9. *Nous* generally defined as the pre-Socratic attribute of Thought, defined materially, not conceived as *psyche (mind/soul)*; *Physis* generally defined as the pre-Socratic attribute of extension, governed by laws of physics (see Anaxagoras).
10. *Pure Immanence* (Deleuze's concept) which I think of as Being immanent to life itself, not anything more than a life lived, not Being understood as immanent to a subject.

11. The contrast of infinities here is positioned as presupposed notion: on one hand, that we must be able to *count* infinity (as it seems in the case in Badiou); or on the other hand, that it is *non-denumerable* (as appears to be the case in Spinoza and Deleuze). My view agrees with Deleuze and Spinoza, insofar as infinity is essentially qualitative and not quantitative, i.e., non-denumerable.
12. The attributes are infinite, equal in power, and distinct expressions of God's essence. There can be no asymmetry at the level of the attributes due to the necessity of God's perfect nature and causal power. Additionally, attributes are equal in quality, and infinite, so there can be no asymmetry, coupling, or quantitative inequality as Badiou claims.
13. Infinite mode is understood as the Infinite Intellect in the attribute of Thought and motion-and-rest in the attribute of Extension. There may be more, but we can only conceive of two attributes: Thought and Extension. I am not discussing the immediate-mediate duality here.
14. One-to-one correspondence, i.e., isomorphism, or a structural mapping between binary systems.
15. Modal Expression for Deleuze seems to be the affirmation of the power of finite being through the expression of the infinite.
16. (Proof at E1P30): A true idea must agree with its object [*ideatum*]…that which is contained in the intellect as an object of thought must necessarily exist in nature…; (Second Corollary at E1P32): will and intellect bear the same relationship to God's nature as motion-and-rest and, absolutely, as all natural phenomena that must be determined by God to exist and act in a definite way (Spinoza 2006, pp. 20-21).

Bibliography

Badiou, Alain. (2005) *Being and Event* (New York: Continuum)

Badiou, Alain. (2015) *Theoretical Writings* (New York: Bloomsbury Academic)

Badiou, Alain. (2022) *Badiou by Badiou* (Stanford: Stanford University Press)

Craig, Patrick. (2013) "*Absoluta Cogitatio*: Badiou, Deleuze, and the Equality of Powers in Spinoza," *Epoché 18(1)*, pp. 227-246

De Beistegui, Miguel. (2012) *Immanence – Deleuze and Philosophy*, pp. 105-159 (Edinburgh: Edinburgh University Press)

Deleuze, Gilles. (1988) *Spinoza, Practical Philosophy* (San Francisco: City Lights Books)

Deleuze, Gilles. (1990) *Expressionism in Philosophy: Spinoza*. 1st ed. (New York: Zone Books)

Deleuze, Gilles, et al. (1994a) *What Is Philosophy?* (New York: Columbia University Press)

Deleuze, Gilles. (1994b) *Difference and Repetition* (New York: Columbia University Press)

Deleuze, Gilles. (1995) *Negotiations* (New York: Columbia University Press)

Duffy, Simon. (2004) "The Logic of Expression in Deleuze's Expressionism in Philosophy: Spinoza: A Strategy of Engagement," *International Journal of Philosophical Studies 12(1)*, pp. 47-60

Lautman, Albert. (2011) *Mathematics, Ideas and the Physical Real* (New York: Continuum)

Kant, Immanuel. (1991) "The Contest of Faculties," in: H.S. Riess (Ed), *Kant: Political Writings*, pp. 176-190 (Cambridge: University Press)

Roffe, Jon. (2007) "The Errant Name: Badiou and Deleuze on Individuation, Causality and Infinite Modes in Spinoza" *Continental Philosophy Review 40*, pp. 389-406

Shein, Noa. (2015) "Causation and Determinate Existence of Finite Modes in Spinoza," *Archiv für Geschichte der Philosophie 97(3)*, pp. 334-357

Spinoza, Baruch, Michael L. Morgan, and Samuel Shirley. (2006) *The Essential Spinoza*. 1st ed. (Cambridge: Hackett Publishing Company)

CRITIQUE OF THE COMMUNITARIAN THEORY OF AESTHETIC VALUE

Theo Morgan-Arnold

INTRODUCTION

This article responds to the communitarian theory of aesthetic value as presented in Nick Riggle's forthcoming paper, "Aesthetic Value and the Practice of Aesthetic Valuing." It will examine what a communitarian theory must capture in order to be a successful theory of aesthetic value. It will then provide an example which demonstrates that Riggle's formulation of communitarianism fails to capture this. Finally, possible modifications to the communitarian theory will be provided and critically analyzed. The only plausible interpretation of aesthetic communitarianism seems to be empty, it does not tell us anything about what aesthetic value is.

INTERPRETATIONS OF COMMUNITARIANISM

The communitarian theory of aesthetic value claims that aesthetic value is just whatever is worthy of engagement in the social participatory practice of aesthetic valuing (Riggle 2024a forthcoming, Riggle 2024b forthcoming). Here is an analogous individualistic theory of aesthetic value: aesthetic value is just whatever is worthy of engagement by a person. There are three ways we can interpret "a person" in the individualistic theory. First, we could say that an object is aesthetically valuable just in case it is considered worthy of engagement by any one person. On this relativistic interpretation, if one person finds children banging on pans with spoons worthy of engagement, then children banging on pans with spoons would be aesthetically valuable. This theory says that almost everything has aesthetic value; this clearly is not what people mean when they claim that an object has aesthetic value. Another interpretation is that it is specifically one person who determines what is worthy of engagement, for example: An object is aesthetically valuable just in case Joe Biden finds it worthy of engagement. This sort of imperialistic theory seems absurd as well. When someone says that an object has aesthetic value, they don't mean themselves, or another individual, are the sole arbiter of what is aesthetically valuable. The third way to interpret this theory is to say: an object is aesthetically valuable just in case an appropriate aesthetic valuer finds it worthy of engagement. This third theoretical interpretation seems more plausible than the

relativistic and imperialistic interpretations. However, it seems empty: the phrase “the appropriate” does all of the work, but does not tell us anything about what constitutes an appropriate aesthetic valuer. In order to make the communitarian theory plausible, we must interpret it in such a way that we avoid both extremes, and hit a middle which actually tells us something non-trivial about what aesthetic value is.

One way which we can read “the” is “at least one aesthetic community.” We will call this interpretation the relativist interpretation. This sort of communitarian theory would claim that so long as any two individuals act in such a way to sufficiently be called an aesthetic community, anything which they deem worthy of engagement would have aesthetic value. Imagine someone accidentally left a pen mark on a classroom wall. Two students see this pen mark and think that it is a piece of minimalist graffiti, they both say that this is an excellent work of art. On the relativist interpretation of the communitarian theory, this would be sufficient for the pen mark to have aesthetic value. According to any interpretation of the communitarian theory, an isolated individual's attention to and judgment of this pen mark would not be sufficient to constitute aesthetic value since aesthetic value is defined in terms of aesthetic community.

When people make claims about the value of a particular piece of art, it seems intuitive that they take themselves to be making more than just the base relativistic claim, “I like this.” For example if one makes a claim, “the Mona Lisa is a masterpiece” they seem to be saying something more than simply, “I like this painting.” If we were to ask any number of people what they mean when they say that the Mona Lisa is a masterpiece, they would likely reference facts about the painting which they believe make it valuable. Most people would not merely say, “I like the way I feel when I look at it.” When people refer to enjoyment while expressing evaluative judgements they would likely explain the pleasant feeling through reference to things in the painting which make it aesthetically valuable. When people make claims about aesthetic value, they are making claims which they believe are at least semi-objective.

In order to see the implausibility of the relativist communitarian theory, imagine a very bad example of an individualist theory of aesthetic value: Any aesthetic object which an individual finds worthy of engagement has aesthetic value. According to this theory, an object has aesthetic value if anyone finds it worthy of engagement. These conditions for aesthetic value seem far too liberal: people think that there are things which are aesthetically valuable and things which are not, and that there is at least some fact of the matter as to which things are aesthetically valuable. This

definition is clearly lacking some level of objectivity which people tend to think is present in aesthetic value. The relativistic communitarian position occupies a similarly dissatisfying position to this individualistic theory. Our theoretical definition of aesthetic value ought to reflect at least some objectivity about aesthetic value since the thing which "aesthetic value" ordinarily means is at least semi-objective.

There is another interpretation of aesthetic community which offers us the greater objectivity we desire. We can interpret "the" as meaning "our" aesthetic community. Let us call this the imperialist interpretation. We can draw the line for our aesthetic community at varying levels of specificity, for example a broad definition might be "post-renaissance western aesthetic community," a more specific one might be "2020s United States aesthetic community." We could imagine that these two aesthetic communities are governed by different value sets. Of course the latter would fit within the value set of the former since it is a component of it. According to this interpretation, what is aesthetically valuable is just what is worthy of engagement in *our* aesthetic community. Since our aesthetic community has a specific set of governing values, these governing values are the objective basis for our aesthetic judgements. For instance, we can now say that the Mona Lisa has far greater aesthetic value than the pen mark on the wall because in the post-renaissance western aesthetic community, works which display mastery of a medium are considered more aesthetically valuable than works which do not. This interpretation gives us the objectivity which the relativist interpretation left us wanting, but seems to take things too far.

Imagine a Martian community which has a long practice of making fourth dimensional "paintings." These "paintings" are invisible to us yet space sociologists tell us that Martian practices surrounding these artifacts are quite similar to human practices of pondering paintings in galleries. It seems like ordinary language users would be able to say something like "Martian fourth dimensional paintings have aesthetic value." People could and likely would truthfully say this, despite the fact that no one in *our* aesthetic community finds them worthy of engagement. In fact, since humans are limited to three dimensional aesthetic experiences, no human will ever be able to have an aesthetic experience of a martian fourth dimensional painting, and will therefore never be able to form an aesthetic judgment of one. A human could correctly attribute aesthetic value to some artifact which their community does not find worthy of engagement. When a human does this, the underlying meaning of what they are saying is something like "Martian fourth dimensional paintings are aesthetically valuable for those who can perceive them." The imperialist

interpretation takes the objectivity so far that we cannot account for the value of aesthetic objects which our community does not deem worthy of engagement. A desirable communitarian theory would be able to account for the correct attribution of aesthetic value of Martian fourth dimensional paintings despite the impossibility of engagement for the person correctly attributing this.

OBJECTION TO RIGGLE'S COMMUNITARIANISM

Both the imperialist and relativist interpretations of the aesthetic community are implausible. Riggle does not advocate for either of these interpretations and defines aesthetic community in a way which tries to satisfy the desire for an appropriate mean between the two insufficient interpretations.

Riggle defines aesthetic community in the following way: "Aesthetic community exists between two or more people when, only when, and because their ways of exercising their capacities for discretionary valuing and volitional openness are mutually supportive. Aesthetic community brings these capacities, and the goods their exercise produces, to fruition and is an end in itself" (Riggle 2024a forthcoming, p. 19). Since we have determined that the imperialist and relativist interpretations of "the" aesthetic community are theoretically undesirable, we will interpret Riggle as saying something analogous to "an appropriate" aesthetic community. An appropriate aesthetic community is an aesthetic community which meets the above definition, i.e., discretionary valuing and volitional openness deployed in tandem in a mutually supportive matter. His approach to defining the aesthetic community seems promising for capturing the middle ground between objectivity and relativity.

Riggle defines the values of volitional openness and discretionary valuing: Volitional openness "is the ability to discount or temper our normal modes of action and reaction and act from an open perspective that is directly engaged with and immediately responsive to one's environment" (Riggle 2024a forthcoming, p. 21). "When we engage in discretionary valuing, we choose to devote our time, attention, care, and concern to the thing of our choice. And we take a certain pride in, even identify with, the results" (Riggle 2024a forthcoming, p. 20).

We can imagine our engagement and attention as water flowing through a valve. Volitional openness acts like turning the valve on our aesthetic attention to the left, we let our attention to objects flow more openly when we are volitionally open. Volitional openness is a value which allows us to be more open to new aesthetic experiences. Discretionary valuing turns the valve on

our aesthetic attention to the right, it closes off our attention to certain things. When the valve is appropriately tuned, we are open to new aesthetic experiences, but not so open that we are limited to only superficial engagement with every aesthetic object we encounter. It is also not so closed that we only have deep engagement with a few aesthetic objects of the sort we are already familiar with.

The limiting factor which picks out an appropriate aesthetic community in Riggle's theory is the requirement of aesthetic communities to have a mutually supportive exercise of discretionary valuing and volitional openness. We will call the thing which Riggle's definition of aesthetic community picks out AC*, and leave AC to more broadly capture instances which ordinary language users might call aesthetic community.

Let us consider a pre-renaissance European aesthetic landscape. The church was the primary commissioner of art in the Middle Ages. Art was almost exclusively used to depict biblical events and portray divinity as the church saw appropriate. People would encounter this art as part of their religious practice; the mostly illiterate populous would learn biblical stories from these depictions. It would be incredibly difficult and likely impossible to find AC*. In this context, people were not deploying discretionary valuing and volitional openness in a mutually supportive way here. In fact, discretionary valuing would have been actively repressed, and volitional openness to non-Christian aesthetic objects would have been punishable. Nonetheless, there were aesthetically valuable works present at this time. Since the communitarian defines aesthetic value in terms of aesthetic community, and AC* does not pick out any community at the time where we would say there is aesthetic value, they are pushed to either:

A) Claim that these objects were in fact not aesthetically valuable,

B) Amend AC* to make it relativistic enough to capture the aesthetic community present at the time, or

C) Claim that these objects were valuable at the time they were made because an appropriate aesthetic community would find them worthy of engagement.

Route A seems implausible: these works were commissioned, maintained over centuries, displayed in places which indicated their importance. If we look at how people interacted with these works, they clearly treated them as if they were valuable. Since these are aesthetic objects, and people treated them as valuable, it seems implausible to make the claim that these artworks lacked aesthetic value just because the thing AC* picks out was not present. One might claim that in fact, these paintings had religious value rather than aesthetic value. This claim is also untenable; there are all sorts of religious rituals and objects which have religious value, but religious value is not mutually exclusive with other types of value. For example, the bible has literary and religious value, a sermon has oratory and religious value, likewise a religious painting has religious and aesthetic value. It is the case that there is a thing which AC* picks out in the world today, this is plain to see if we sit in an art museum coffee shop and listen to people's conversations. There are medieval paintings which are considered aesthetically valuable which are hanging in museums now. It would be quite odd to make the claim that these paintings lacked aesthetic value when they were created, and only came to have aesthetic value when in the post-renaissance world AC* emerged. When people say that these medieval paintings are aesthetically valuable, they take themselves to be making some claim about the painting itself. We might be in a community that is in a better position to recognize the aesthetic value of these paintings since members of our community mutually deploy discretionary valuing and volitional openness. However, being in a better position to recognize the aesthetic value of an object, and being the sole basis for the existence of aesthetic value in an object are radically different things. The former seems like a modest and plausible claim, the latter seems entirely untenable. Since the communitarian could not plausibly deny the existence of aesthetic value in these paintings prior to AC*, they must abandon route A.

In taking route B, the communitarian might claim that discretionary valuing and volitional openness are not necessary components of AC*. There might be aesthetic communities which repress the exercise of these faculties, or which only minimally deploy them. Riggle seems to endorse this view, "it is conceivable that some groups realize aesthetic community in other ways, via other characteristic actions or by downplaying some of those emphasized here. For example, maybe some groups frown upon self-expression or restrict sharing for cultural, political, or religious reasons. They can nonetheless have culturally significant ways of realizing aesthetic community" (Riggle 2024 forthcoming, p. 25). Route B is amending Riggle's AC* so that it more

appropriately captures aesthetic community. The issue with AC* is that it seems to be too imperialistic; we would need an amendment which widens the scope of AC*. If the communitarian makes discretionary valuing and volitional openness merely optional but unnecessary for AC*, they would fall into relativism. One solution might be to say that AC* is just a community which has the appropriate amounts of discretionary valuing and volitional openness. This solution does not seem plausible because in order for the theoretical definition of aesthetic community to be informative, it needs to tell us something about what "appropriate" means. We already know how to appropriately use the term "aesthetic community" in our language to pick out things in the world, our theoretical definition is not clarifying what it is appropriate to apply the term to, but trying to clarify the grounds for appropriate application. To theoretically define aesthetic community as whatever someone might appropriately call aesthetic community leaves the theory rather empty. Riggle's AC* does not map on the appropriate application of the term "aesthetic community" in English, a tenable communitarian theory would need to give an account of aesthetic community which plausibly captures the meaning of the term. Route B does not need to be abandoned, but it requires a reworking of the definition of aesthetic community.

If the communitarian takes route C, they might tell a story about the convergence of the aesthetic community. For example: Throughout history there have been oppressive regimes and heavy religious influence restricting the free development of aesthetic community, however we have reached a point in history where the values which constitute aesthetic community have become clear to us. We can retroactively see what was aesthetically valuable historically even though AC* was not present during this period in history. Since we have converged on appropriate aesthetic community in the present, aesthetic works that our aesthetic community finds worthy of engagement can be retroactively recognized as aesthetically valuable. The belief that we have converged on some more true form of aesthetic community in the present seems untenable. Value claims which base objectivity in convergence have been made in every era. For instance, one could imagine a medieval art enjoyer claiming, "the Greeks thought that they understood what was aesthetically valuable in their nude portrayals of human form, but that was only because they had yet to converge on the Christian grounds of aesthetic value, Piety, modesty, devotion." The medieval Christian could just say that Greeks didn't recognize the vulgarity of the nude human form because they had yet to converge on Christian values, they could use the dominance of Christianity in their present to justify this claim. It is not clear what, if any, non-arbitrary grounds

could be provided for an argument of the superiority of one form of aesthetic community over another. If a communitarian endorses a definition of aesthetic community which is non-relativistic, they must motivate the criteria for distinguishing between superior and inferior forms of aesthetic community.

The other primary issue with route C, is that the theoretical explanation of aesthetic value seems to be empty. Consider the following theory: An object is aesthetically valuable just in case we can say of it truthfully, "this has aesthetic value." This theory accurately captures most or all cases of aesthetic value, yet it tells us nothing about what aesthetic value *actually is*, it is accurate because it is empty. We can see that the individualistic theory which we offered earlier in the paper is similarly empty: an object is aesthetically valuable just in case an appropriate aesthetic valuer finds it worthy of engagement.

CONCLUSION

This article has shown the three types of interpretations for a communitarian theory of aesthetic value. It has addressed why relativistic and imperialistic interpretations are untenable and identified the plausible middle ground. It has then demonstrated that this plausible middle ground is unable to account for aesthetic value in all circumstances. If the communitarian is to account for aesthetic value in these circumstances, their theory lapses into either relativity, imperialism, or emptiness, which are all untenable positions.

BIBLIOGRAPHY

Riggle, Nick. (2022) "Toward a Communitarian Theory of Aesthetic Value," *The Journal of Aesthetics and Art Criticism*, Volume 80, Issue 1, pp 16–30

Riggle, Nick. (2024b forthcoming) "Autonomy and Aesthetic Valuing," *Philosophy and Phenomenological Research*

Riggle, Nick. (2024a forthcoming) "Aesthetic Value and the Practice of Aesthetic Valuing," *The Philosophical Review*

Fear and Loathing in the State of Surveillance: A Philosophical Inquiry into the Ethics of Mass Surveillance and the Necessity to Further Develop a Right to Privacy

Nate Miller

Introduction

Government surveillance and the practice of intelligence collection and analysis in the name of national security present a major problem for the supposed right to privacy. Mass surveillance has not only become more prevalent in society today, but more accepted as well. With cameras on every street corner, microphones in our pockets, our businesses, and our homes that listen in on every conversation, and major corporations eager to sell our data to the highest bidder, we must consider what privacy means to us today. The major theory in support of this surveillance is the Just Intelligence Theory (JIT) proposed by Quinlan (2007) in *Intelligence and National Security*. It is also necessary to consider the adaptations made to JIT by Miller (2021) in *Social Epistemology*. In contrast — and arguing against the indiscriminate use of mass surveillance — are the theories of privacy as a matter of control and privacy as a matter of access proposed by Macnish (2018) in the *Journal of Applied Philosophy*. These theories will form the basis of my analysis, but I will also incorporate many articles on privacy, security, and intelligence throughout to support my arguments and claims. In this article, I will explicate how the concepts of JIT and privacy as access both fall short of meeting the challenge mass surveillance poses to the populace. I will argue that privacy is a fundamental moral right and should be defined as a matter of control. I will demonstrate that these distinctions are essential to maintaining human dignity and freedom, and that acting on privacy as defined in this manner will contribute to individual, communal, and societal flourishing.

The extent to which government surveillance and data collection is being practiced in the United States is truly troubling. Journalist Glenn Greenwald has done substantial research on this topic and published many seminal articles detailing the surveillance we are subject to. This information is specifically laid out in the work Greenwald has produced and contributed to based on the information and documents he received and discussed with National Security Agency (NSA) whistleblower Edward Snowden. For example, Greenwald identifies “the Verizon

document" as one of the most damaging documents exposed (Public Broadcasting Service 2014, p. 25). The Verizon document proved that the United States government was collecting the phone records of all Americans legally under contract with Verizon via a secret court order. This document revealed that "the NSA surveillance system is not directed at very bad people…it's directed at the American citizenry and other citizenries around the world, indiscriminately, in bulk" (Public Broadcasting Service 2014, pp. 8, 25).

Furthermore, while this information was collected as metadata, Greenwald explains that the metadata can infringe on privacy more significantly than audio records alone could. Specifically, he argues that "you can be incredibly invasive into the most intimate and private aspect of people's lives by collecting metadata," explaining an example of a woman contacting an abortion clinic to support his point (Public Broadcasting Service 2014, p. 26). In this example, if you were listening in on the call, you would just hear a person schedule an appointment at a generic-sounding hospital clinic. In contrast, by gathering and analyzing the metadata, you would know exactly what the person is doing, whom they are doing it with, and what their network of friends and associates looks like (Public Broadcasting Service 2014, p. 26). A government with a history of invasive surveillance practices, a reckless disregard for any configuration to the right of privacy, and vast amounts of new sources and data available to it all combine to create frightening future possibilities. However, before discussing how defining a right to privacy would help curtail these affronts, it is necessary to understand how mass surveillance is justified by its supporters and implementers.

In Section 1 of this article, I will articulate how JIT and the added measures proposed to it by Miller argue for the continued and enhanced use of intelligence as ethical. I will then evaluate this conception of JIT according to Miller and show it to be lacking in requisite respect for ethical concerns about the right to privacy. In Section 2, I will demonstrate the difference between privacy as access and privacy as control, illustrate the weak and unprotected nature of privacy as access, and argue that Macnish mischaracterizes privacy as control. In Section 3, I will make the case that privacy is a fundamental human right and must be defined clearly and effectively to ensure its status in the modern surveillance state. Furthermore, I will defend my argument that privacy as a fundamental right should be defined according to the "privacy as control" account explicated in section two. Finally, I will illustrate how our contemporary understanding of privacy in the age of

mass government surveillance requires this re-defining of privacy and enables the free development and flourishing of individuals in accordance with their natural rights.

SECTION 1: JUST INTELLIGENCE THEORY AND ITS MISGIVINGS

Just Intelligence Theory (JIT) is derived by former politician and intelligence theorist Michael Quinlan through the analogous use of Just War Theory (Quinlan 2007, p. 1). As in Just War Theory, JIT is divided into two categories: "what we might …say about the conditions under which it may properly be engaged in — *jus ad intelligentiam*, as it were — and about the limitations that ought to be observed in carrying it out even when it is properly undertaken — *jus in intelligentia*" (ibid., p. 3). In understanding this theory, it is also necessary to understand what qualifies as intelligence. It is important to note that in the arena of JIT, the term intelligence is used to refer to what has been introduced previously in this article as surveillance; in the following descriptions, the two will be used interchangeably.

In defining intelligence, Quinlan wants to focus more on intelligence work and how the term can be defined through the specific acts of surveillance and information gathering that commonly take place. For example, he argues that much of intelligence is "collating and analysing material that is available to anyone with the resources and inclination to gather it" and that no moral problem arises from this (ibid., p. 4). Instead, for Quinlan, the moral ambiguity arises when considering eavesdropping, which includes "picking up communications not meant to be heard…[and] watching activities which the actor would have preferred to keep unobserved" (ibid., p. 4). This is where the paradigm of JIT becomes a necessary tool to adjudicate governmental and organizational choices and actions. However, it is dismissive to state that no moral problem arises from gathering and analyzing individuals' personal information without their knowledge or consent, regardless of the accessibility of the information. This perspective will be discussed further in sections 2 and 3. At this point, it seems that privacy rights and control of one's information are paramount. Still, Quinlan argues that when JIT is used to distinguish cases of entitlement and constraint, it can be an effective model for the use of surveillance.

Entitlement corresponds to the criteria of *jus ad intelligentiam* – or when it is legitimate to collect information. Constraint corresponds to the criteria of *jus in intelligentia* – or the methods of collection that must never be used in surveillance (ibid., p. 6). This information constructs the necessary background of JIT required to consider its application in determining the permissible

and impermissible practices of surveillance. JIT, as laid out above and in its traditional form, purports that it can provide adequate guardrails for surveillance when the parameters of moral permissibility are filled out in more detail. However, when considering the implications of the right to privacy as a matter of control from sections two and three of this article and applying them to the context of JIT, it will be clear that JIT cannot result in a more fundamental and ethical use of surveillance. In contrast, philosopher Seumas Miller argues that to make JIT a practical theory for ethical surveillance, you must incorporate the principles of discrimination, necessity, proportionality, and reciprocity; moreover, these considerations salvage the theory as a paradigm for understanding the use of surveillance.

In defining intelligence, Miller importantly notes the distinction that "definitions can be purely descriptive (defining what intelligence *is*) or they can be normative (defining what intelligence *ought to be*)" and then proceeds to make his case for intelligence according to both standards (Miller 2021, p. 212). Miller describes intelligence simply as information or data and ascribes it as a teleological (or purpose/goal-oriented) function that varies based on different institutions and their objectives (ibid., p. 213). In terms of "national security intelligence," the purpose is tautological, i.e., mass surveillance is conducted to protect the nation and its people. The normative theory constitutive to surveillance is JIT, and Miller asserts that through an analysis of JIT, a set of conditions can be understood under which "national security intelligence collection, analysis and dissemination is…morally justified" (ibid., p. 214). Discrimination, necessity, proportionality, and reciprocity make up the *jus in intelligentia* conditions under which surveillance is morally justifiable. Discrimination requires that innocent persons ought not to be harmed or have their rights violated (ibid., p. 218). However, Miller also states that this principle is applied quite permissively in intelligence work – so that while it *ought* not to be the case, it often *is* the case, as seen in the Verizon documents exposed by Greenwald and Snowden. The principle of necessity refers to the level of harm caused by intelligence activity. More specifically, if harm is a necessary byproduct of intelligence gathering, the least harmful means to achieve that end should be chosen. This conception of necessity is inherent in that it is a moral principle, not merely a rational one, because of the "moral quality of the ends in play" (ibid., pp. 219-220). The proportionality principle "ascribes positive moral weight to the ends realized…and negative moral weight to the harms caused" (ibid., p. 223). Appropriate application of the principle of proportionality requires considering the most positive and negative moral consequences.

Reciprocity is the principle Miller hopes would bring about an equilibrium state of affairs among nation-states by ensuring each responds equally to the intelligence activities of the others. What Miller fails to account for is the problematic use governments make of this means-end reasoning to create a veil of moral justification for their encroachment on the rights of law-abiding citizens domestically and globally.

These principles that Miller argues refine JIT are understood as a means-end relationship, especially those of discrimination, necessity, and proportionality. To extrapolate further, the moral consequences and ethical lines crossed in the gathering and use of surveillance (the means) are weighed against the surveillance results and those moral consequences (the ends). Philosopher Adam Moore sums up and exemplifies a problematic case of this means-end rationality by organizing his "nothing to hide" argument as follows (Moore 2011, pp. 145-146):

> (P1) When two fundamental interests conflict, we should adopt a balancing strategy, determine which interest is more compelling, and then sacrifice the lesser interest for the greater. If it is generally true that one sort of interest is more fundamental than another, we are warranted in adopting specific policies that seek to trade the lesser interest for the greater interest.
> (P2) In the conflict between privacy and security, it is almost always the case that security interests are weightier than privacy interests. The privacy intrusions related to data mining or National Security Agency (NSA) surveillance are not as weighty as our security interests in stopping terrorism. These sorts of privacy intrusions are more of a nuisance than a harm.
> (C1) So it follows that we should sacrifice privacy in these cases and perhaps adopt policies that allow privacy intrusions for security reasons.

The argument presented by Moore here logically orients the premises and conclusions described by Miller in his development of JIT. Miller argues for the same considerations outlined in premise one, just through the lens of the principles of discrimination, necessity, and proportionality. JIT, as laid out above in the context of the aiding principles Miller argues it requires and the logical format this argument takes, is not a practical solution to implementing and maintaining the moral use of surveillance tools. The misrepresentation of the weightiness of security interests in the name of national security tips the scales unfairly. As Greenwald documented, governments are willing to

assign little to no weight to privacy rights. These facts combine to show how arguments like the one Moore explicates and arguments for JIT and its variations (such as Miller's) fall short of meeting the challenge mass surveillance poses to privacy. Furthermore, I believe the conditions Miller finds necessary to apply to JIT make it problematic. Specifically, Miller is too loose in drawing moral lines around the principles of necessity, proportionality, and discrimination that he favors over JIT (Miller 2021, pp. 218, 224). Miller makes problematic use of means-end reasoning, often arguing that the means justify the ends, regardless of the scale and scope of the moral bounds they cross, especially in relation to privacy. While the argument provided validly interprets moral dilemmas regarding surveillance and privacy, I do not believe it is a just solution to the problem. In section 2, I will address some specific examples from Macnish's 2018 article that demonstrate variations of the argument presented here and offer my counter-analysis.

Section 2: Privacy as Access, the Problems Associated with Considering Privacy as Access, and the Misrepresentation of Privacy as Control

The arguments for JIT put forth by Quinlan and Miller — and re-organized by Moore — make the case for the perspectives that led to the government actions described in the Snowden files. JIT lays the groundwork on state actors to base their justifications for privacy violations. It enables the mirage of constructing weighty moral ends (i.e., national security) to justify whatever means are necessary to accomplish the goals of an elite few. As employees and directors of intelligence agencies are not elected officials, they must refer to and give weight to the rights of citizens demanded by both legal and natural law. Macnish instead responds to the revelations of the Snowden files and attempts to construct a definition of privacy that can explain the inherent violations that are harmful (Macnish 2018, p. 417). To achieve this goal, Macnish focuses on the distinction between defining privacy as access and defining privacy as control. Macnish argues that privacy should be defined as a matter of access rather than control and provides several examples he believes illustrate his point and defend his argument.

In this section, I will argue against this view and instead make the case that privacy as control is a more accurate way people regard a right to privacy. I will use Macnish's examples to point out the fault in his logic and demonstrate why privacy as control offers a more realistic description of how people orient a loss of privacy. My argument will focus on consent and the perspective that a loss of control over your information (and personal space, as is shown in the

example of the Peeping Tom) forces you to act differently than you would if you could enjoy the oxymoronic freedom of control. However, it is important to note that it is the privacy-violating act that occurs when you do not have control of information about yourself that causes your actions, and not the other way around, as Macnish argues. The comfort of knowing your privacy is in your own hands and can only be loosened or restricted by you, and your consent is essential to understanding privacy as control.

Defining privacy as access means that a loss of privacy only occurs when an individual or organization accesses one's information. In contrast, privacy as control holds that a loss of control over one's information constitutes a loss of privacy (Macnish 2018, p. 417). Julie Inness, one of the leading scholars on privacy as control, defines the concept by stating that privacy is "a variety of freedom, a freedom that functions by granting the individual control over the division between the public and the private with respect to certain aspects of her life" (Macnish 2018, p. 419, Inness 1992, p. 42). However, Macnish concludes that "a loss of control over one's information does not entail a violation of privacy, (but) it does entail a violation of the right's that privacy protects" (Macnish 2018, p. 419). He believes this is an effective position because it allows him to demonstrate the problematic nature of mass surveillance in accordance with a definition of privacy organizations behind this activity would accept.

Macnish provides several scenarios and thought experiments that he argues strengthen the view of taking privacy as access to be an effective definition. For example, Macnish considers a scenario in which you forget your diary at a coffee shop. You return to the coffee shop to retrieve the diary and find it sitting on a different table in front of a stranger. The stranger insists that they respect and appreciate privacy and, therefore, did not open the diary but collected it upon noticing you left without it with the goal of returning it to you. Macnish argues that no violation of privacy has occurred here, according to the access account, because the diary was not opened. Furthermore, he equates this to the government collecting your data but not accessing it and argues for the same conclusion (ibid., p. 420). What Macnish fails to recognize is that in the case of the government collecting your data, it is never deleted or returned to you as the diary is. It is compiled *ad infinitum* until it becomes useful to them somehow. Understanding this requires that you define privacy in a way that articulates this violation. Therefore, it is necessary to define privacy according to the control account. Defining privacy in this way is necessary because it guards against considerations like those proposed by Miller and Macnish. It keeps the power of one's personal information under

the control of the individual and grants this right to all people. This also seems to be the way we generally think about privacy. When we consider privacy, especially in the context of personal information, it entails that the information is known and available only to ourselves, thus making it private. It is ours to maintain or relinquish control of at our own will. If someone else gains control of this information against or without our knowledge and can access it at any time, this would violate privacy. The next example Macnish brings up helps to further examine the accounts of privacy as access and privacy as control as described and in the context of government surveillance.

Macnish wants to distinguish between what he believes are actual privacy violations and what he considers to be the problematic actions people are forced to take to avoid privacy violations (ibid., pp. 422-423). Macnish argues that in the case of a Peeping Tom looking through your window and an eavesdropper trying to listen in on your conversation, no privacy violation occurs if you avoid the Peeping Tom seeing you or the eavesdropper overhearing you. The significance of this distinction for Macnish is that it notes an important difference in how privacy is defined in the access account versus the control account. In this example and Macnish's argument, the control account of privacy includes an individual's evasive actions to maintain their privacy, e.g., hiding from the Peeping Tom (ibid., p. 423). Again, he believes this is a misplaced definition of privacy because only a threatened violation of privacy has occurred, not an actual violation. However, Macnish concedes that in the Peeping Tom example, at minimum, a violation of "private space" has occurred (ibid., p. 423). Here, the distinction I believe Macnish is trying to make is that of the rights of liberty and security regarding property and the body versus the rights of privacy. However, I would counter that privacy in these contexts, and that of government surveillance cannot be conflated with the rights of liberty and security. In all these examples, the right to privacy refers to the privacy of personal information about oneself. Furthermore, in all these examples, this right to privacy is violated by the party attempting to (or in government surveillance successfully) obtain this information without your consent. What follows from this line of reasoning is that privacy must be defined according to the control account to accurately capture what we really consider to be private. To further defend this account of privacy, it is necessary to explain why privacy is a fundamental human right and, specifically, how control, consent, and freedom influence this understanding.

SECTION 3: THE FUNDAMENTAL RIGHT TO PRIVACY

It is important to define privacy more explicitly for many reasons, the most important of which is that the argument for privacy as a fundamental moral right and requiring control over one's personal information has not been articulated in the face of mass government surveillance. In response to the arguments of Miller and Macnish, it is important because establishing a clear right to privacy as a legal standard and human right would enable more robust accountability for organizations engaged in mass surveillance. Furthermore, a human right violation holds stronger ethical weight than an action that is simply considered amoral. A description of this difference and its normative implications provides a strong response to the arguments proposed by the proponents of JIT and Macnish's conception of privacy as access. Philosophers Mark Alfino and G. Randolph Mayes agree that privacy is a fundamental moral right. However, they argue that even so, "the right to privacy is not sufficient to justify the claim that individuals are entitled to control over their personal information" (Alfino & Mayes 2003, p. 2). Privacy is "fundamental" in that it cannot be explained by referring to other rights, such as those of security or liberty as described before. Privacy is "moral" in that its ends are considered good, ends everyone would agree to. Privacy is a "right" in that it is something people are entitled to "rather than a mere good that any rational person might strive to achieve" (ibid., p. 2). As explained in section 2, I agree that privacy is a fundamental moral right, a distinguished right not captured by rights of security of the person or property, as critics of this view often argue. In contrast to Alfino and Mayes, I believe defining privacy in this way does lend itself as a fundamental right to the notion of control.

The fundamental moral right of privacy protects an individual's personal information and the good that comes with knowing others do not have access to that information without your consent. The good described here is the opportunity for self-exploration and growth, protection from unwanted harassment, and freedom from coercion and restriction. This is a significant difference from the case Alfino and Mayes present, where influencing a person to act differently is not considered coercion or something the right to privacy protects you from (ibid., pp. 13-14). The problem with Alfino and Mayes' perspective here is that they believe this is the case because no violation of privacy occurred beforehand which caused these consequences. For example, they argue that the most obvious feature of their analysis is that "a violation of privacy consists in a form of interference with a person's activities" and that situations such as intruding into someone's personal space "is not reasonably construed as a form of coercion" because you are still free to

choose to act in any way you please (ibid., pp. 13-14). In the context of government surveillance, this analysis does not seem to apply so straightforwardly.

The contention that privacy and coercion are mutually exclusive and can pertain simultaneously is incompatible. A violation of privacy will always directly impact those whose privacy has been violated and will affect their behavior. This resulting behavior is not due to these individuals freely choosing to be affected by the preceding violation in that way but is an act limited and tied to that event. The example of government surveillance helps to identify why this is not just a *post hoc* analysis. Government surveillance is inherently an exercise of power. The purpose of this surveillance is "national security," the prevention of crime and terrorism, and an essential practice in achieving that end is forcing people to act differently under this umbrella of mass surveillance. When you know the government is collecting the metadata of your search history, phone calls, and text messages, applying facial recognition software, and planting cameras on every corner, you are forced to act differently to maintain some semblance of privacy. The strength of the control account of privacy is that it places the power of genuine freedom to choose your actions back into the hands of the people. Alfino and Mayes argue that their analysis resembles John Stuart Mill's principle of liberty but does not reduce to a theory of personal liberty (ibid., pp. 13-14). I think they are correct in stonewalling this reduction but mischaracterize the logic as not applying to a theory of privacy founded in control. For example, in chapter five of *On Liberty*, Mill contends that there are three main objections to government interference, even in support of its citizens. Most relevant is Mill's third objection, which argues that "the most cogent reason for restricting the interference of government, is the great evil of adding unnecessarily to its power" (Mill 1869, p. 312). While power is clearly an important determinant in the principle of liberty, again, it is not mutually exclusive from privacy. The power that control over one's information brings to a fundamental right to privacy does not necessarily reduce it to liberty. Instead, this power prevents invasions of privacy and maintains the necessary component of privacy, which is one's personal information. The benefits of liberty that result are just that, but it is the right of privacy that enables them, not a right of liberty. Philosopher Thi C. Nguyen, referencing Onora O'Neill, argues that transparency is not wholly good and that trust and transparency are often opposed and in conflict (Nguyen 2021, p. 332). Nguyen's article supports my argument that control of our information, i.e., privacy as control, is necessary for maintaining freedom among the populace and trust in the institutions we elect and direct to govern us. The

examples of Nguyen and Mill are important to include as conclusions to my argument because they are representative of the multitude of thinkers who believe privacy is a fundamental right, an inherent good, and something that contributes to a healthy, flourishing society and that these facts are necessary to its construction.

CONCLUSION

In this article, I have argued that privacy is a fundamental moral right and that the control account of privacy offers a practical, necessary definition in the state of surveillance that consumes the contemporary age. I have maintained that the control account of privacy defends a logically constructed and endoxic view of privacy, meets the challenge mass surveillance poses to society and contributes to a worldview endorsed by logic and real-world examples. I have also argued that defining privacy as access refers to an account of privacy that identifies the concept incorrectly. Furthermore, I have shown that the theories of JIT assign the right to privacy improper moral weight and employ faulty means-end reasoning to defend government surveillance. The goal of this article is to re-think how we define privacy and what it means to us in an age where it is persistently under the threat of being diminished to a relic of past human experience.

Bibliography

Alfino, M., & Mayes, G. R. (2003) “Reconstructing the right to privacy,” *Social Theory and Practice* 29, pp. 1-18

Inness, Julie. (1992) *Privacy, Intimacy, and Isolation* (Oxford: Oxford University Press)

Macnish, Kevin. (2018) “Government Surveillance and Why Defining Privacy Matters in a Post-Snowden World,” *Journal of Applied Philosophy* 35, pp. 417-432

Mill, John Stuart. (1869) *On Liberty* (London: Longman, Roberts, & Green Co)

Miller, Seumas. (2021) “Rethinking the Just Intelligence Theory of National Security Intelligence Collection and Analysis: The Principles of Discrimination, Necessity, Proportionality and Reciprocity,” *Social Epistemology* 35, pp. 211-231

Moore, D. Adam. (2011) “Privacy, Security, and Government Surveillance: Wikileaks and the New Accountability,” *Public Affairs Quarterly* 25, pp. 141-156

Nguyen, C. Thi. (2021) “Transparency is Surveillance,” *Philosophy and Phenomenological Research* 105, pp. 331-361

Public Broadcasting Service. (2014) *The frontline interview: Glenn Greenwald – United States of Secrets*. PBS. https://www.pbs.org/wgbh/pages/frontline/government-elections-politics/united-states-of-secrets/the-frontline-interview-glenn-greenwald/

Quinlan, Michael. (2007) “Just Intelligence: Prolegomena to an Ethical Theory,” *Intelligence and National Security* 22, pp. 1-13

A Sympathy for Affective Architecture

Joe Vargas

Introduction

This article attempts to present a case in which Baruch Spinoza's overlooked philosophy can provide insights into addressing design problems within contemporary architecture. Design within contemporary architecture, especially since the mid-twentieth century, has been especially influenced by philosophical inquiry to designate architectural problems.[1] Because it may not be immediately obvious as to what is the connection between architecture and philosophy, this connection is through our existence in reality. Architecture involves designing the built environment to construct our reality, while philosophy involves questions about what our reality is. This connection is the seat of a formation of the field known as architectural theory. And often, ontologies are particularly important within architectural theory, simply because they account for the existence of the material in objects that manifest as structures.

Architecture theory in the mid-twentieth century focused on consequences of capitalism negatively affecting the design of social spaces.[2] Contemporary architecture theories broadened the focus onto other problems like global climate change. The philosophers and architects Peg Rawes and Gökhan Kodalak have recently made great strides in interpreting Spinoza's philosophy as a nascent architectural theory, meanwhile, Lars Spuybroek developed a comparable architectural theory through his interpretation of John Ruskin's philosophy. This article will be a comparative analysis of Ruskin's sympathy and Spinoza's affects as the method to bootstrap a more tenable architectural theory of Spinoza.

In section 1, I will provide a general basis for philosophy as an informer of architectural theory to address architectural problems. In section 2, I will explore architectural elements expressible through a Gothic ontology. There will be an aesthetic focus on John Ruskin's concept of sympathy. In section 3, I will give a brief account of Spinoza's ontology. Lastly, in section 4, I will compare the Spinoza's ontology with the Gothic's ontology with to justify an underlying geometric connection to stipulate comparable architectural theories. There will be a focus on Spinoza's concept of affect.

Section 1: Architecture Theory

This section should introduce some of the necessary context to step into the mindset of a contemporary architectural theorist. Architecture is more than the appearance of structure. Instead, there are very real problems to consider such as structural soundness, protection from the elements, energy usage, climate change, material sourcing, the construction process, specific client and city requirements, social and economic-class issues, budget requirements, and the overall lifecycle of the structure's materials and its surrounding environmental impact. Solving all these problems in architecture starts within the architectural design process (La Roche p. 31).

It's the primary competency of the architects to become an export of this architectural design process. The realized design of every built structure is a direct reflection of its said design process. The success of a design is understood by how the material forms of each built structure solves its architectural problems. As a simple example, a basic pavilion would be required to provide its inhabitants shade from sunlight, to protect from rain, and to remain stable from inflicting forces. Assuming the pavilion had been constructed exactly as it was designed, any failure to address any of these three requirements reflects as some type of failure within its design process.

However, this does not necessitate a conclusion that if a design is successful then it is the ideal design. In addition, this is not to conclude that the best design of all possible successful designs, by all knowable metrics, implies the existence of an ideal design process where this ideal design process will universally produce ideal designs. Instead, this notion of an ideal architectural design process alludes to an important divide within the history of architecture and to what architects believe they have been actually trying to achieve through the production of forms. Architect Farshid Moussavi explains this division as two different systems to approach design depending on its notion of a design principle.

> Throughout the history of architecture the development of built forms has unfolded within two systems that could be broadly described as top-down, in which a single principle determined the relationship between an ideal whole and its constituent parts, and bottom-up, where a single principle determined a system of parts that were repeated to produce the whole (Moussavi 2009, p. 27).

The top-down system is a system of the ideal architectural design. There is a fixed idea what the design should resemble as a whole and its parts should resemble the whole. For example with the ancient Greeks, mathematics exists outside our realm of existence, and it was instead from a realm of eternality. Thus, mathematics gave them access to an ideal through transcendence in design. The design adheres to the model of a fixed ideal. They applied the mathematics in their designs through geometric ratios and symmetry.[3] The materials were crafted as the parts of architectural elements that formed the geometric design. Together the architectural elements, such as columns, walls, roofs, doors, and stairs, constitute the parts forming the whole that resembles the mathematical ratios and symmetry. This ideal whole-from-parts relationship is how the ancient Greeks strived for beauty in their designs.

On the other hand, the bottom-up system is not a system of an ideal design, but rather, it emerges through a design process that adapts over iterations. Today, architects largely lean towards this iterative approach to design. This design process allows for an incorporation of viable feedback during the process. Each iteration within the process presents the opportunities to learn from feedback and apply it into the design. Feedback of suggestions or needs might be given from other architects, engineers, future patrons, or the architect's own critique. Basic feedback of suggestion may include examples such as adjusting the location of an apartment complex's entrance to be safer by locating it farther from the street. This may cause a big shuffle of all the floor plans starting from the ground level upwards. Another feedback of need is increasing the number of units in the apartment complex while appropriately balancing a tradeoff of spacious units. It's the architect's responsibility to transform the collection of feedback into their decision making on the design. Thus, during each iteration, the adaptions improve a solution linked to a set of architectural problems. Each adaptation will occur in tandem with other adaptations. This process will iterate multiple times until the completion of the design. The completion of the process is successful when everyone's needs have been adequately addressed throughout the design process.

This iterative approach in the bottom-up system does not adapt itself to a fixed and final ideal of an architectural form. Instead, the system's single principle is a geometry that is an anexact protogeometry. It's a type of geometry that has a capacity of becoming into existence through an innate pattern during its growth. This growth is similar to a snowflake's growth which is partially influenced by environmental factors yet also partially influenced by a self-assembly from its

protogeometry of the hexagon. Because the hexagon is the inherent geometry within snowflakes, it's the protogeometry. As a bottom-up type of system, the snowflake is not forced to necessitate a resemblance to any ideal geometry in the whole. To the contrary, if it's true that every snowflake is unique, then all snowflakes exist differently while still sharing the same protogeometry within its parts. This is due to the inherent capacity for adaptability within the snowflake's hexagon. The adaptability allows the protogeometry to act the base unit in the bottom-up system. During the bottom-up system's geometric formation, the single principle is the protogeometry that causes the generation. The generating process adapts and composes the expressed parts of a hexagon protogeometry into a whole of the snowflake structure. The result is a somewhat repeating, yet uniquely coherent pattern of the snowflake. Through this analogy, this immanent generation of the parts into whole is the same as the protogeometry within architectural elements that constitutes as the parts of a coherent, yet unique architecture whole (Moussavi 2009, p. 31). There is still a big difference between a snowflake and a formally comparable hexagonal architecture. Although the snowflake is just as formally coherent, the architecture also doubles the coherency of its geometric form to function as the solution to architectural problems.

How are these problems exactly addressed through geometric forms? Consider a simple problem of needing an inner courtyard in the design of a solid circular building. One functional solution is to carve out a rectangular space which forms the inner courtyard of the circular building. However, another solution that is more formally coherent and just as functional is to carve out a circular space which forms the inner courtyard.

While a formally coherent and functional space is a good solution, it's unfortunately not the complete answer as to why there should be an inner courtyard included to the design in the first place. Perhaps, it was given as part of feedback, but that still doesn't necessitate its priority among many other architectural problems that may rigidly conflict. One example may be retaining the needed solution in the design for a self-contained area with a large square footage to function as a lecture hall. If the lecture hall overlaps with space of the proposed inner courtyard, then that's a design problem with no flawless solution. Ultimately, recognizing the architectural problems worth addressing is an important function of an architectural theory. And although the top-down and bottom-up systems are two different systems, they are representative of a larger pattern among all architectural theories.

One might imagine a design scenario for an academic institution. The circular building with the crowded learning spaces of large lecture halls reflects one set of priorities because it holds more students in total. Or, the circular building with the calming space of an inner courtyard but with smaller lecture halls reflects a different set of priorities because it holds fewer students in total. This is simply to say that there isn't a singular or fixed architectural theory that is universally applicable.

Thus, how and which architectural theory is applied within the design process becomes pivotal, as prioritizing architectural problems is prioritizing values. In other words, the decision in favor of an architectural theory becomes an ethical decision because a tenable theory frames the question of whose and what needs are more valuable to address. This notion of a tenable theory is still in addition to the generation of geometric forms becoming the built environment which expresses these prioritized values through the construction of our reality.

While there may be various ways to distinguish between architectural theories, architect David Leatherbarrow describes architectural theories as dialects among “bodies of ideas prepared for advance and defense”. Over the past decades, theorists and practitioners alike have engaged in a tug-of-war between different paradigms, each vying for dominance and relevance (David Leatherbarrow 2023, p. xv). In a similar tradition, the remainder of this article is in the tug-of-war of architectural theory between a Ruskin approach and a Spinozian approach.

SECTION 2: GOTHIC SYMPATHY IN ARCHITECTURAL ELEMENTS

Architect and theorist Lars Spuybroek developed an architectural theory from an immanent ontology of the Gothic by interpreting philosopher John Ruskin’s concept of sympathy. Sympathy is the basis for his singular design principle in a bottom-up system. Although this architectural theory is similar to the hexagonal snowflake because it is bottom-up, the design principle of sympathy doesn't stipulate a distinct protogeometry acting as the basic unit. This is unlike the relationship between the protogeometric hexagon and the structure of the snowflake. Instead of protogeometry, sympathy is fundamentally a principle of immanent relations. The geometries of Gothic structures emerges through sympathy. Sympathy is one type of immanent relation, so one should examine exactly what is meant by this immanent relation in the Gothic ontology.

In the Gothic ontology, there are entities that exist through an intertwining or entangling. The entanglement forms a relation. The entities twist together like the ribs of a Gothic structure,

encroaching into each other's space leaving no gaps in between. The Gothic ribs somewhat resemble the delicate twigs of a bird's nest that tightly pack together. Because the twigs form a thing of the nest, each twig becomes more than the mere physicality of dead tree-matter. In other words, a twig's identity isn't derived from intrinsic properties limited to be a part of a tree, and instead its identity is from a relational dance in its lifecycle of interactions. The identity was once distinctly part of a tree, and it later adapted as distinctly part of a nest (Spuybroek 2016, pp. 44, 48-49). In this example, a sympathy is within the relational dance among the twigs that immanently causes the emergence of the nest. This is possible because within the inherent capacity of twig-matter is to relate in a nest-like manner. This appears strikingly similar to the growth of a snowflake because of the inherent capacity in the protogeometry of a hexagon to adapt and compose, but the twig has an emphasis of a relation instead of protogeometry.

To push this emphasis further, one must first clarify how Gothic ribs compare to the aforementioned twigs. Of course, they are still similar because both neither abide to a fixed ideal of identity that retains a notion of some essence. If this wasn't the case, the essence of a twig would be limited to be a tree part because the tree-ness of a tree determines all of its parts to be aligned within the tree's identity, cause, or purpose. Instead, it is necessary that both the twigs and the ribs have an unfixed identity as an adaptable part.

The difference between them is the degree in which their motions as relational parts cause and persist the emergence of the whole structure. This is because Gothic ribs have a great emphasis on motion while twigs in the nest are simply set tightly packed together. The Gothic ribs remain in a constant motion for expansion that claims all space, leaving no gaps among the ribs. As if it were a nest of constantly moving twigs, this motion forms a symphony of interaction between each thing of action. It's a "thing-action" (ibid., p. 44). Thus, one can define a thing-action as the immanent and sympathetic rib in constant motion that constitutes the Gothic structure.

To make this clearer, one can consider a rib to be a part of a whole in a wild plant, and the plant is set in a crowded space among other wild plants where each is vying for survival. There's an inherent push within the wild plant for its growth. This wild plant pushes up against the other plants in an effort to fight them for more space or be forced into an involution from its overgrowth, as there are no gaps available to be filled. On the contrary to the wild plant, the tamed plant doesn't push, as it much as it prevents being pushed in order to survive.

The lack of gaps is a fundamental feature to the fabric of matter in the Gothic ontology. The boundary between the ribs of the plants is where the plants push, creating a byproduct of life. It's where new life can form and old life can die. In this way, there's a larger dance among the vying plants that constitutes a sort of beauty. It's a "savage beauty of life". The beauty occurs between the wild and tamed plants in this type of relational dance. Because the wild must fill and tamed must retain, beauty is in the interplay. Through this aesthetic, all created things become like plants, including houses, paintings, and even countries (ibid., p. 156). It's as if the whole world of existence is a single, all-encompassing structure made up of different plants that are held together in a symphony of interlocking gestures of the Gothic ribs. Thus, sympathy is present where there's a direct conduit or connection for the "immediate flux of life" at the boundaries of the ribs, and beauty is present where there's a balancing interplay between different plants against their different sympathies. In an effort to reassert thing-actions in order to introduce group-things, the Gothic rib and the plants, which are in constant motion with a shared synchrony through sympathy, are now both defined as thing-actions. Therefore, the groups of interrelating plants, which are in constant motion with beauty at their boundaries, are then defined as "group-things". A group-thing constitutes the structure of space as an ecology of plants.

How does this all-encompassing structure of thing-actions and group-things manifest as geometries such that they can be understood architecturally? As a bottom-up system, the geometries in Gothic architecture emerge up from the different motions of sympathy into a motion of beauty of the space. For the sake of this argument, one can distinguish "public" and "private" as two ends in a continuum that is a property of space. A public space can be an interior space like a large lecture hall, or an exterior space like a courtyard. A private space is normally an interior space, like a bedroom. Because a private space is normally an interior space, more architectural elements such as walls, a floor, a roof, a door, and a window are used as the insulating geometry for increased privacy. Still, a public space also will contain architectural elements, but allows for the possible interactions of its inhabitants.

The public-private division among the spaces in and about a structure has the starting point in its maximal ornamentation. This is namely through the thing-action's geometry that emerges from the sympathy of Gothic ribs in formation of the plants. The growth of the plant as an ornamentation has moments of differentiation. It's where the wild crosses its boundaries into the tamed, and during the move of the wild into the tamed emerges as a new.

The growth and differentiation of the Gothic ornamentation is similarly expressed at a larger scale of architectural spaces at their boundaries, the transitional spaces. In between the larger, more public spaces furls out the smaller, more private spaces such as rest areas like nooks, restrooms, or even storage space. Spuybroek argues that the architectural growth in the Gothic is a design without a designer. In other words, the spaces do not fit towards an ideal geometry, and instead, they immanently react at moments during the growth of life. In the Gothic, the growth creates differentiations of ornamentations that emerge up into differentiations of architectural elements, which then the different architectural elements emerge up into differentiations of architectural spaces. It's a chain of relations that moves sympathy from the small scale to the large scale, yet it also imbues a Gothic beauty through the coherency of differences within the design.

This chain of relations frames a geometric and ecological relationship of a Gothic structure and its environment. Because the structure comes to exist with the geometry of an existing environment, they come together, forming the built environment. In other words, a structure becomes a part in our reality, not only through its direct existence to us alone, but it also indirectly becomes a part indirectly through the surrounding impact it has on other things and the environmental realities. These other things generally encompass other people, animals, plants, structures, infrastructure, etc. Meanwhile, the environmental realties are more persistent over time because they encompass either climate related forces or difficult to capture matter. These include sunlight, heat and cold, wind, waterfall, water runoff, and pollution. Through a network with other things and environmental realties, we are part of a group-thing such that we do not share a mutually immediate experience with everything.

Consider a set of geometry in a contemporary Gothic structure that adapts more according to an environmental reality of heat instead of another thing like the existing electrical grid. For a higher quality of life for inhabitants, heat in a structure can be dealt with actively through air-conditioning or it can be dealt with passively through built-in ventilation. Air-conditioning has been the hallmark contemporary appliance for interior temperature control. Unfortunately, an air-conditioning appliance introduces two glaring problems that negatively impacts its surroundings. The first and most obvious problem is a required usage of electrical energy for the air-conditioner. And often, this appliance operates using the electricity sourced from a network of maintained infrastructure from the area's electrical grid. The second and less known is the cooling-agent inside air-conditioner. For any air-conditioner to cool, it requires the use and production of butane gas, a

volatile organic compound. Volatile organic compounds are toxins which slowly contribute to irreversible damage to bodily organs and potentially neurological damage whenever someone is overexposed to a leakage. In addition, the other pollutants during production of volatile organic compounds can harm the factory workers that produced the compounds (Bergman 2012, pp. 93-96).

This forceful or imposed method of handling issues within these environmental realities is known as an active design strategy (ibid., p. 67). The negative consequences from this type of design strategy can be quantified more generally by a carbon footprint. The footprint reflects the carbon impact onto the environment over an object's life cycle from its production, energy usage, and death.

Now instead, one can consider how a Gothic structure may adapts its geometric growth with a passive design strategy. It would prioritize a geometrically vertical growth that allows for a high chimney in order to deal with heat. The high chimneys found in Gothic cathedrals, like the Notre Dame Cathedral or the Cathedral of Cologne, solve a common architectural problem known as the "stack effect". This is where hot and unventilated air rises to the ceiling of a room, then it fills the room downwards. This air is considered polluted. It is both uncomfortable and harmful for inhabitants. On the other hand, a high chimney, especially paired with a higher ceiling and open windows, pulls the hot air upwards and outwards using the thermal pressure and wind pressure already present in the environment. Thus, the temperature and air quality of the Gothic room can be solved through a passive design strategy (Li & Qian 2016, pp. 61-63).

To address a counter concern, there may be more materials incorporated with the construction of a high chimney, and therefore, there's an increased carbon footprint from the production and transportation of the chimney's material. While there is some truth for the concern after the initial construction, it's always important to weigh the design approaches against the overall carbon footprints, and this is in addition to accounting for a prevention of harmful exposures. In this case of this structure with a high chimney, there's the reduced energy usage relative to air-conditioning and less pollution from the reduced production of volatile organic compounds found in air-conditioning. Therefore, there's a lower projected carbon footprint for its life span. Through the Gothic geometry enacting with a passive design strategy, it's then more arguable that a Gothic structure will sustain a more resilient beauty within its larger environmental network, its group-thing.

SECTION 3: SPINOZA'S ONTOLOGY OF IMMANENCE

In the *Ethics*, Spinoza encapsulated an immanent and monistic ontology. His system of concepts relies on three fundamental concepts: substance, attributes, and modalities. It's important to notice an easy misinterpretation of these words. While Spinoza mimics these terms for these concepts from preceding philosophers, his usage is distinct. For example, Spinoza gives fundamentally different meanings of these terms that contrast meanings given by Descartes. In one case where the difference is clearly apparent, Spinoza argues that his conception of God exists as an extended thing, meaning in physical space, and this ultimately underlies his ontology with a conception of immanence, as given in the *Ethics* at Part 2, Proposition 2 (Spinoza 1677, p. 115). This perspective is fundamentally incompatible with a Cartesian ontology because Descartes's conception of God can't exist in extension.

Spinoza's conception of God derives from his concept of substance. According to Spinoza, a substance must be an infinite and eternal essence[4] that constitutes all individual things and their differences.[5] Then, he is able to determine that there is only the need for a singular substance because an infinite essence has no bounds for its ability to constitute all things. This one substance, because it constitutes everything, is Nature. And if God is also infinite and eternal, then God is the same substance as Nature (Kodalak 2018, p. 90). They have exactly the same meaning because the words also are interchangeable in meaning. This forms the basis of his ontology as a monism with immanence.

If God and Nature are the same, he thereby attacks the Judeo-Christian constraints for a divine, supernatural, and transcendent creator, who created a separate existence of our natural realm that includes our existence in extension. Instead, there is no transcendent dimension governed by sovereign and a divine entity that is separate from us (Kodalak 2020, p. 37). There is only the immanent interactions within the eternal cosmos as God or Nature. For this reason, his conception of an immanent God must, at minimum, diverge from an anthropomorphic conception of a transcendent God (ibid., p. 36).[6] If an immanent God constitutes the full reality of all existing things beyond humans, then one must abandon the underlying sense of a hierarchical division that's implicit in transcendence along with the claim that only humans are created in the direct image of a transcendent God.

This hierarchical division in transcendence has the types of existence that splits between the ideal essence of a human and whatever individuals conform to it. In other words, the human individuals imperfectly conform to this transcendent essence of a human, so the expressions of human individuals are beneath this type of a transcendent existence. However, if one does not split existence as a hierarchy of levels, then one would conceive the different types of existences as a heterarchy within one level or plane. This one level of heterarchy is the one substantial plane of existence, or specifically, it's a heterogenous plane of immanence (Kodalak 2018, pp. 90, 94). Here, this one level is Nature or God, so there does not exist a supernatural level above it. One type of expressed existence is not necessarily above or precedes another (ibid., p. 91). Thus, different people cannot be more or less conforming to a human essence than another; different things, whether human, animal, plant, or architectural structure, is not more or less substantial than another. To the contrary, they all have a comparable capacity for existence and interaction, because they are constituted through the same one substance of God or Nature. This is Spinoza's basis for how immanence differs from transcendence.

One can't directly understand the substance of God or Nature in its totality, either through perception or intellect. God or Nature is unbounded in its expressions and is expressed through infinitely distinct ways. Two of these expressive ways are extension and thought. Although there are infinitely other ways to express because God or Nature is infinite, these two ways are the only ways which people can understand the existence of things because people are similarly finite through these expressions and that limits understanding. Spinoza defines attributes as these ways for the expressions of the substance into things, and these things which are expressed through the attributes are called modalities.[7] Humans,[8] animals, plants, structures, and everything else that we can understand through perception or intellect are individuals composed of modalities. While all these modalities are different, they originate from and share the same, one substance of God or Nature. However, unlike God or Nature, modalities are finite and temporal because they existentially are created and destroyed.[9] They are not instantiated as the infinite substance which is eternal and without a cause other than from itself,[10] but rather, modalities occur as the causal effects of the substance through its attributes (ibid., p. 91). Therefore, the conceptualization of Spinoza's ontology is the claim where the one and singular substance, God or Nature, potentiates its infinite and immanent capabilities through its attributes into the modalities across its heterogenous plane of an immanent existence (ibid., p. 90).

SECTION 4: RIBS AND MODALITIES TO SYMPATHY AND AFFECTS

The concept of modalities is an important entry point for the ontological comparison with the Gothic ontology because modalities have similar behaviors to the Gothic ribs as immanent things. For example, Spinoza argues that modalities are infinitesimally simple, yet they also compound into larger bodies, allowing them to fill in the gaps of existence, such as the modalities through extension, which has no void.[11] Just as the Gothic ribs fill up all of the gapless space, so do modalities fill up all of extension.

As another example, one can recall that Gothic ribs want to be in a constant motion, and this motion may ultimately result in a change of identity of its thing-action. The Gothic ontology constrains the identity of thing-actions to be temporal because the underlying Gothic temporality emerges through motion. Modalities, including the bodies which modalities constitute, are similarly temporal as described through Spinoza's concept of motion-and-rest. Motion-and-rest provides the basis for Spinoza's account on identity. The speed of modality's motion-and-rest is what distinguishes a body, as simply a collection of modalities, to be instead identified as an individual, which can now be differentiated from other bodies or individuals.[12] Furthermore, it's within the nature of an individual to persist its identity indefinitely (Kodalak 2020, pp. 85-86).[13] This is akin to self-perseverance within a thing is the concept of conatus (Kodalak 2018, p. 97). If the conatus is within the individuals constituted by finite modalities, then its individuals have the capacity to be destroyed by external causations[14] applied onto these modalities. Here, new individuals can emerge from the changes to modalities of the old individuals. The identity of the old individual disappears when a new configuration of replaced modalities sufficiently and relationally differ than to the previous configuration, according to changes with their motions-and-rests.

One also can recall the similar explanations for identity within the Gothic ontology. The emergence of the nest's identity, which occurs through the configuration of the twigs, will later dissipate when the twigs sufficiently reconfigure away from a nest-like form. This is the emergence of the Gothic thing-action, as it is like the emergence of the Spinozian individual body. An individual's conatus is also similar to the persistence of identity as described by the pushes of the Gothic plants. The tamed plant pushes for survival, retaining its identity, and the wild plant pushes for growth, either expanding its identity or potentiating a new identity through causing a nearby

reconfiguration. The conatus of the individual desires to retain its identity, thus it similarly fights for its individual's survival. In this fight, it also fights against the occurrence of new identities. At the loss of the conatus's fight, a modal transformation occurs, whether expressed through extension or thought, but never crossing attributes, from the same and singular substance. Then emerges a new individual with a new conatus.

Kodalak also interprets an individual's motion-and-rest as "an immediate infinite modality harboring infinitesimal multiplicities that constitute differential relations through which individual modalities come to be and differ from each other", for he later concludes that "individual modalities constitute ever-changing manners and moods of the cosmos, [meaning we] are all different rhythms, unique ways of life" (Kodalak 2018, p. 95). Relations and rhythms are as fundamental to Spinoza's ontology as they are to the Gothic. However, their ontological accounts for identity between individuals and thing-actions differ yet overlap between the comparison Spinoza's affect and Ruskin's sympathy.

Fundamentally, they are both different yet similar because of the way individuals experience other individuals. Affects encompass feelings, or emotions, and generally bodily sensations because the affect serves as a driving force behind all action and experience within individuals. An affect can be joyous or sorrowful, which is not to be confused to be taken as the literal emotion, but it should be taken rather through a positive sense of empowerment, or affirmation, and a negative sense of disempowerment within the individual to act or to experience events. An interpreted account of affect through Gilles Deleuze's ontological system, which is deeply informed by Spinoza's ontology, explains the function of affect as the following:

> The mechanism that operates the Deleuzian ontological movement is *affect*. Following Spinoza, a body is not a fixed entity but a dynamic relationship, whose limits are subject to change depending on what it encounters affectually. When two bodies meet, they are either compatible or incompatible, they either compose a new relationship or decompose each other. (Frichot 2013, p. 252)

The affect is experienced because of an external cause where there's a bridge with another individual. This bridge is the concept of affection because it connects individuals or bodies. Despite its name, an affection is never to be confused with an affect. As the bridge, an affection

forms a reciprocal, or bidirectional, relation between individuals that experience affects, within a shared environment (Kodalak 2018, p. 98). An affection doesn't necessitate both individuals experience the same affect of joy or sorrow in ordered for their affects to be affirmed.

Affects with affections among modalities and individuals encompass the functionality of sympathy among ribs and thing-actions. Because sympathy necessitates the same feeling to form a compatible relation, Spuybroek sympathy claims to emphasize a deeper and more synchronous connection than affects. The feelings with sympathy resonate as a mirror and interpenetrate neighboring thing-actions, while affects are too mechanical for a strong resonance in congruency of a network, or group-thing (Spuybroek 2016, pp. 173-174). On the other hand, the overall goal here is to encompass a near enough ontological superset of Spinoza over Ruskin, in order to justify bootstrapping a more tenable architectural theory for Spinoza. Being too mechanical is an irrelevant claim against affect and affection applying to architecture. The only requirement is that a Spinozian architectural theory can permit a strong resonance in a network of individuals which demonstrates their overlapping ontological functionalities.

As an example network, there's a painter using a paintbrush on canvas in a room. It forms a network of four individuals in resonance with each other that enacts the painting. This network is also the becoming of another and larger individual that is non-human and architectural (Kodalak 2018, pp. 104). This also means this type of network encompasses the functionality that closely resembles the Gothic beauty. Because the Gothic beauty emerges from the balance of the Gothic ribs up into the new identity of its group-thing, the conatus of the network constitutes from the tension of the painter-paintbrush-canvas-room and its temporary stabilization of an identity. Eventually, the tensions will erupt into varying speeds and directions of motion known as differential velocities. These differential velocities ripple throughout the network, causing a disruption into new configurations and patterns. Thus, the differential velocities characterize the emergence of the new identities from the previously individuated network (ibid., p. 95). Similarly, the differential velocity emerges like the Gothic beauty that emerges. When beauty emerges, it's from changes in the minute motions of the Gothic ribs through the network of the group-thing. In this moment of change and disruption, the networks of both kinds can expand in their geometric growths.

CONCLUSION

First, we saw that contemporary architects follow an iterative based architectural design process to produce the geometries inherent in materials to solve real architectural problems, architectural theorists inform their design processes with philosophy. Second, an interpretation of Ruskin via the Gothic ontology demonstrated how an architectural theory based on sympathy can construct a Gothic geometry. Lastly, Spinoza's immanent ontology was provided to inform an ontological comparison with the Gothic ontology, giving some credence to a Gothic geometry within a Spinozian architectural theory.[15]

NOTES

1. Architecture theorists have relied on, namely and non-exhaustively, the works of Edmund Husserl, Martin Heidegger, Merleau-Ponty, Jacques Derrida, Michel Foucault, and Gilles Deleuze (Kodalak 2018 p. 105).
2. Architectural theorist Manfredo Tafuri acted as the pivotal figure for widening the discipline's perspective by involving a critique of politics, literature, and philosophy. In examination of the relationship between Modern architecture and history, he became concerned with the lasting socioeconomic effects represented in the subjecting stories expressed through the 'avant-garde' architecture that was being designed. Drawing inspiration from various sources like Hegel, Nietzsche, Benjamin, Adorno, and Foucault, he critiqued the design attempts of the architects to address this crisis. He considered the designs futile, because attempts he believed that the core and unaddressed issue in architecture was a problematic relation between architecture and capitalism (Haddad 2023 p. 6).
3. Ancient Greek architects incorporated the Golden number in their architecture because it resembled beauty. It was applied a part of the proportional design that used the Golden section (Moussavi 2009, p. 27).
4. In the *Ethics* at Part 1, Proposition 11, "God-in other words a substance consisting of infinite attributes, each of which expresses eternal and infinite essence-necessarily exists" (Spinoza 1677, p. 82)
5. In the *Ethics* at Part I, Proposition 15 "Whatever exists exists in God, and nothing can exist or be conceived without God" (Spinoza 1677, p. 86).

6. As a larger relevant topic with this article, it's whether or not and the degree to which anthropomorphic conceptualizing should occur within architectural theory. It is often a highly controversial topic as it relates to both subverting perception and subverting values. Rawes interprets the underlying motivational instance against this centering humanity through Spinoza's conception of substance, "Substance/Nature immanently constitute a plenitude of realities in different modes and scales: from the divine to the common, and from the non-anthropological entity or environment to the scale of the singular being. Substance's productive power therefore constructs nature's diversity, its potential for change (Natura Naturans), and its various modes of existence (Natura Naturata)" (Rawes 2012, p. 63). She later expands how might some anthropomorphic thinking entail some progress with sustainability efforts, "Spinoza's natural geometry therefore produces common-place ecologies and common lives. Common notions are ecologies or 'life-places' of diverse human subjectivities and relations. Such biodiversity is not just a neutral or value-free materialism, but has the politics of equality at its core. This communal immanence does not inevitably result in unfettered anthropomorphic progress or unethical infinity, but accords with feminist philosophers 'ethical biological, cultural and social ecological thinking for all" (ibid., p. 66).
7. The term "modality" is often used synonymously with the term "mode".
8. In the *Ethics* at Part 2, Postulate 1, "The human body is composed of very many individual of a diverse nature, each of which is highly composite" (Spinoza 1677, p. 130).
9. In the *Ethics* at Part 1, Proposition 28, "Every particular thing, or, any thing which is finite and has a determinate existence, cannot exist or be determined to be operate unless it is determined to existence and operation by another cause, which is also finite and has a determinate existence; and again, the latter cause also cannot exist or be determined to operation unless it is determined to existence and operation by another cause, which is also finite and has a determinate existence, and so on to infinite" (Spinoza 1677, p. 98).
10. In the *Ethics* at Part 1, Proposition 7, "It belongs to the nature of substance to exist," and its Demonstration, "A substance cannot be produced by something else; therefore, it will be the cause of itself. That is (by definition 1), its essence necessarily involves existence, or, it belongs to its nature to exist. QED" (Spinoza 1677, p. 78).

11. In the *Ethics* at Part 1, Proposition 15, Scholium, “Since, therefore, no vacuum exists in Nature, but all its parts must agree in such a way that a vacuum does not exist, it follows also that they cannot be really distinguished; that is, that corporeal substance, in so far as it is substance, cannot be divided [as finite parts]” (Spinoza 1677, p. 89).
12. In the *Ethics* at Part 2, Proposition 13, Lemma 1, “Bodies are distinguished from each other in respect of motion and rest, speed and slowness, and not in respect to substance” (Spinoza 1677, p. 126).
13. In the *Ethics* at Part 3, Proposition 6, “Each thing, in so far as it is in itself, endeavors to persevere in its being,” and later at Proposition 8, “The endeavor by which each thing strives to persevere in its being involves, not a finite, but an indefinite, time” (Spinoza 1677 p. 171).
14. In the *Ethics* at Part 3, Proposition 4, “Nothing can be destroyed except by an external cause” (Spinoza 1677, p. 170).
15. If this article were to expand, it would explore into the provocative ideas given by Rawes. She asserts that Spinoza's content and presentation of the *Ethics* makes it inherently architectural. It deeply connects geometry and psychology. “Like the inherent relational connection between the mind and the body, Spinoza also considers mathematical and aesthetic reasoning to be related activities” (Rawes 2018, p. 4). “Spinoza compellingly shows that these geometric relations arise because of the connection between our mental and physical experiences. Geometry is not a mathematical procedure or constitution of ideal or disembodied forms which are unknowable to the individual (i.e. the issue of the ‘divided line ’in Plato’s theory of ideal forms, in which reason is only always/ever divine and unknowable, and thereby entirely distinct and exclusive of the sensible, yet inconsistent and frail, human subject). Instead, the *Ethics* presents geometry as an inherently relational form of reasoning (and a distinctly human power)” (Rawes 2018). She also holds the view that we may have need to rid the use of the computers in design due to limitations with embedded in its normative representations in form. Instead, we can use the *Ethics* as a guide for developing an alternative method for producing non-normative geometry. This is because modern CAD software through parametric tools and techniques have tainted by a false claim for a “natural/value-free universalism” in their

geometric productions. Rawes disagrees with Spuybroek's appeal to use digital tools to generate geometry, Gothic or otherwise (Rawes 2018, pp. 61, 68).

BIBLIOGRAPHY

Bergman, David. (2012) *Sustainable Design: A Critical Guide* (Princeton Architectural Press)

Frichot, Hélène. (2013) "Abstract Care," in: *Deleuze and Architecture,* pp. X-X (Edinburgh University Press)

Haddad, Elie G. (2023)" What ever Became of Architectural Theory?" *The Contested Territory of Architectural Theory*, pp. 3-15 (Routledge)

Kodalak, Gökhan. (2018)" Spinoza, Heterarchical Ontology, and Affective Architecture," in: *Spinoza's Philosophy of Ratio,* pp. 89-107 (Edinburgh University Press)

Kodalak, Gökhan. (2020) *Spinoza and Architecture* (Cornell University)

La Roche, Pablo. (2017) *Carbon-Neutral: Architectural Design* (CRC Press)

Leatherbarrow, David. (2023)" Fighting for life," in: *The Contested Territory of Architectural Theory*, pp. xv-xxii (Routledge)

Rawes, Peg. (2012b) "Spinoza's Geometric Ecologies," *Interstices: Journal of Architecture and Related Arts* 13, pp. 60-70 (Auckland University)

Rawes, Peg. (2018) "Aesthetic Geometries of Life," *Textural Practice* vol. 33, iss. 15, pp. 787-802

Moussavi, Farshind. (2009). "The Function of Form", in: *The Function of Form* (Actar and Harvard University Graduate School of Design).

Spinoza, Baruch. (1677) *Ethics*, edited and translated by G.H.R. Parkinson (2000), (Oxford University Press)

Spuybroek, Lars. (2016) *The Sympathy of Things: Ruskin and the Ecology of Design* (Bloomsbury Visual Arts)

Yang, Li and Fene Qian. (2016) "Research on Green Systems of the Historical Buildings Protection," *Proceedings of the 2016 International Conference on Energy, Power and Electrical Engineering*, pp. 61-65 (Atlantic Press)

"HEAVEN" IS REAL

Edie Markovich

Modal realism is the view that there exist concrete possible worlds which are of the same kind as our "actual" world. Introduced by David Lewis in the late 80s, this quite radical stance sparked commotion within the philosophical community. In this paper, I shall begin by grappling with Lewis's conception of reality, then contrast it with other views, and hopefully come to a conclusion that falls more in alignment with my intuitions. However, before venturing further, I find it only necessary to answer a fundamental question: what is a world? Does it solely encompass our beloved planet Earth, or does that term extend further into the vast universe? One way to look at a world is as a *limit* or a situation that includes a collection of physical objects that follow some sort of recognizable structure. This limit can be more or less inclusive. According to this view, the entirety of the physical universe as it expands out into space and time would constitute the "actual" world. However, I don't appreciate the groundedness of that description, so let us take the term 'world' to refer to the realm which we inhabit and investigate. Now, on an arguably more important note, I don't find it at all reasonable to limit our conception of the universe, or worldly realm, to that which is *physical* because in doing so, we are closing the door to several very interesting dilemmas. Where dilemma here should not be seen as possessing a negative connotation, but rather as a gateway leading one toward progress and discovery. Thus, let the world we speak of include the physical universe as well as the *metaphysical* universe, which includes abstract objects such as numbers and other concepts that humans generally muse about. Sometimes it is said that these abstract objects exist in Platonic heaven, an abstract realm comprised of ideal concepts, or Platonic Ideas. This side of the universe transcends space and time. Which makes one wonder, is it really real?

A modal realist would tell you that the actual world is just a notable example of a possible world, one that doesn't differ in kind from all other possible worlds. And that the actual world is only actual because it is *our* world, but that doesn't remotely come in the way of stripping other worlds of their claim to be truly and concretely real. In these assertions I identify a certain degree of disrespect regarding what it means to be actual and what it means to be real. Reality is actuality,

whereas reality and possibility are radically different notions, and it would be utter sacrilege to equate the two, which is what modal realists seem to be doing.

As we know, one of the fundamental claims made by modal realists is that all possible worlds are of the same kind, namely the kind that constitutes our actual world. This denies what I take the word 'actual' to entail. Possibility, impossibility, and necessity deal with what can, cannot, or what has to be actual. All varying cases of probability that something is in existence. So, it doesn't make sense to say that the actual world is actual because of the fact that it is *our* world, and that all possible worlds are just as real. When I say that something is not actual, I am not saying that it doesn't belong to our world or that the object I speak of isn't here. I mean to say that it is *nowhere*. It all comes down to the idea that possibility is just a degree of actuality in the sense that there exists potential to be actual. And what is actual most certainly does not fall under the "lesser" category of what is merely possible.

Basic possible world semantics is supposed to help clarify the notions of possibility and necessity, which, used colloquially, seem to be relatively straightforward. It's not the case that something is possible because it does not contradict the laws of physics that govern our world, reality. Rather, possibility is just truth in some possible world and necessity is truth in *every* possible world. However, it doesn't feel right to say that it is *necessarily* true that WW2 would have occurred in any possible world. The fact that WW2 occurred in our actual world is not dependent on anything itself necessarily true. All the events leading up to WW2 could have been otherwise, so the culminating consequences of those events could have been otherwise just the same. It seems to be the case that a definition of necessity utilizing possible world semantics refers to a far more specific notion, one of *logical* necessity. Consider the fact that all circles are round. To say that fact is true in every possible world would be entirely acceptable simply due to the fact that I cannot, in any capacity, conceive of a circle that is not round. This is just one example of why modal realism may not be the most appropriate view, because it lends itself to a niche understanding of necessity. One that may not be deemed useful in non-philosophical contexts.

In my mind I envision the actual world living amongst other possible worlds in the same meta-dimension, one that is a step beyond the dimension of the only world we are able to perceive and experience, the so-called actual world. Which could be due to the limitations placed upon us by our own consciousness. But worlds that contain illusory realities of what could have been are not the same kind as the one world containing the baseline: the way things *are*. And any deviation

away from that baseline, however minimal, becomes something that never was, something that evolved from nothingness into mere possibility.

However, charity must be lent to Lewis and all other modal realists. Their point was not to misrepresent or deny the implications of what it means to be actual. Instead, they just wanted to present it in the same light as indexicals. An indexical is a word whose referent varies depending upon its speaker. For me to say that 'I am hungry' varies in meaning as opposed to my friend saying, 'I am hungry', even though we both technically said the same thing. The answer lies in the only indexical in that phrase: 'I'. Because my friend and I are not one and the same instantiation of existence, the word 'I' in the first phrase refers to someone entirely different from the one being referred to in the second phrase. Lewis wants to say that the word 'actual' behaves in the same way as the word 'I' does. So, for me to say that I live in the actual world would just be to say that I live in the world that I belong to. Which strikes me as incoherent. Just to belong to a world does not imply that world is the actual world because the actual world is the only world that ever was and ever will be. Perhaps, sometime in the future, it will mimic a possible world that it differs from in the present moment, but at one particular moment in time it will always be the only world. For everything that is indeed in existence in the present dictates the divide between what could have been possible and what is necessary at time t.

Further outlining the modal realist's muddled picture of reality, they claim that non-actual *possibilia* are just objects that aren't *here*, objects that don't belong to our world. But what it means to be and not be here is where the falsity lies. Because there is nothing *but* here. There is no other sector of reality that includes worlds that contain the possibilities of something that isn't here but could have been. There is only what was in that present moment, the entirety of everything that is here forms the actual in its entirety. This leaves no room for worlds of possible existence, a mere glimpse at what could have been, a dangerous mirage that draws one in with the temptation of its possibility. Salvation from the actual, a necessary predicament bestowed upon all those who belong to its sphere.

If Lewis were right and possible worlds did concretely exist, there would be one notable implication: time travel. Some possible worlds at time t that differ from the actual world at time t can conceivably be in alignment with the actual world at time q. Which means that a person belonging to the actual world at time t can travel to the possible world that will become the actual world at time q, where time q is some time in the future. If the one actual world and the rest of the

infinite possible worlds really did exist in the same meta-dimension and were concretely available to be received, then an individual in the actual world would be able to jump to possible worlds that are true in the future or were true in the past. Time slices of the actual world inserted amongst time slices that were never actualized, only theorized.

Moving forward with a practical example … whether I am a lacrosse player in this world or in any other possible world is accidental according to Lewis and most of the remaining strict or rigid modal realists. In other words, it is not necessary that I am one thing or another in any possible world. Counterparts of my individual self are said to exist. Which means that it must be the case that objects in a possible world, and of course the actual world along with them (because it was said to be just a notable example of a possible world) are their own counterparts. And it is only a matter of chance that this counterpart and everything that defines its state of existence has been actualized. There was nothing *logically* necessary about the fact that I am this way in the actual world. That is simply the possible world that was chosen to actualize itself, almost as if driven by god. But there is also a degree of necessity associated with an object during its permanence through time. Freddie Mercury might have necessarily been of Parsi-Indian descent, but it is only accidental that he decided to form the band, Queen.

So, we've established that on the view of the modal realist, possible worlds do exist and are of the same kind as our "actual" world. A possible world is also an irreducible entity, meaning that it can't be reduced to anything more basic. Also essential to their view is the idea that the term 'actual' is an indexical. There are two more fundamental ideas associated with modal realism. The first is that possible worlds are unified by their spatiotemporal relations, and this is how they distinguish themselves from one another. It follows from this that objects that constitute distinct possible worlds are worldbound, which compels the idea that possible worlds don't overlap. However, at least on the Lewisian conception of modal realism, these spatiotemporal relations are oddly flexible, allowing for the possibility of non-spatial entities such as spirits (which might not be located in physical space, but are at least located in time).

Secondly, possible worlds are causally isolated from one another; there is no spatiotemporal overlap between worlds. In other words, nothing that happens in one possible world has any sort of effect on what happens in another possible world. For example, if a nuclear war were to break out in one possible world at time t, a world that will never be actualized, that has no effect on the actual world. If this were not the case one can only imagine how radical the actual

world could have been. The actions of one of your counterparts that belongs to another possible world would be felt in your world and vice versa. Which would completely change the laws of cause and effect. The actions that we perform as individuals have their own consequences. But imposing the consequences of our other-worldly counterpart on our own, actual existence, would be pure chaos.

To consider all possible worlds just as real as the actual world would be to undermine reality and introduce a dangerous sort of freedom into our thinking, one that would dilute the facts that we must at least partly abide by. The word 'actual' in 'the actual world' is not the same as the word 'I' in 'I am hungry'. In fact, the word 'actual' is something very concrete and not contextually dependent. A distinction must be made between the world that constitutes reality and the possible worlds that constitute possibility, something *not* yet actualized, but can be actualized in the future. Lewis's conception of the word 'actual' to be an indexical straightforwardly begs the question. As we know, if different people say, 'I am hungry', the referent of that 'I' changes depending on who says it. But if different people say, 'the actual world', the referent of 'actual' doesn't change. The only time it would change (and act as an indexical) is if someone from a possible world used it. But then that would assume that possible worlds do indeed concretely exist. So, it doesn't make sense to treat the term 'actual' as context dependent because it's not. Every single time the word 'actual' is used is *within* the actual world. And each time it is used it refers to *one* reality, the only real reality. To claim that 'actual' in 'the actual world' is an indexical, presupposes the existence of a context in which some world is possible, or impossible, or not actual. None of which I believe truly exists apart from present ideations about future possibilities. Which leads us to time …

… an important factor in this whole equation. It seems to be the case that on Lewis's view possible worlds exist at the same time as our actual world. I would disagree, although I would say that possible worlds exist in the future (not the past), as entities that can be actualized. These future possible worlds are causally dependent upon our present actual world and are largely dictated by what occurs in our world. I say not the past because the past is only a recording of actuality. The possibilities have either actualized themselves or turned into logically necessary falsities. Once the present has gone past, the only thing that remains is a shadow of the actual, and a denial of what went from possible to simply false and inherently unactualizable.

As opposed to concrete modal realism, there exists abstractionism. On this view, a possible world serves as the limit of a process. Where this process involves adding to and extending an

initial state of the current world. In other words, it is a *total* way reality could have been, one that encompasses every possibility. The limit of these seemingly limitless possibilities is constrained by consistency. That is, not even the merest detail can be added to a possible world on the abstractionist view without rendering it inconsistent. Which leads us to States of Affairs (SOAs) abstract, intensional entities.

SOAs constitute an ontological category, which is a group of logical concepts used to describe possible things that exist, where logic here is taken to be concerned with understanding the nature of reality and how it is configured. However, SOAs are not defined in terms of possible world semantics. Similar to how propositions function, a State of Affairs can either be actual or not, which is just to say that it either exists or doesn't exist. For example, if I were to say that it's raining outside, that is a State of Affairs that doesn't currently exist, a condition that the concrete form of reality does not currently embody. However, in a few days, that condition has every potential to be actualized, thus transforming into an *actual* State of Affairs. And it is this point that serves as the primary factor that distinguishes modal realists such as Lewis from abstractionists.

Abstractionists treat possible situations, or ways in which reality can be configured, as states or conditions of the physical world as opposed to something necessarily concrete. A State of Affairs is not a structured piece of the physical world, but instead a purely *logical* relation. Thus, abstractionists are generally taken to be less radical because their assertions exist in an abstract realm. They don't posit the existence of other possible worlds that exist in actuality alongside our "actual" world. And while this view may appeal to a wider range of intuitions, it does make fewer substantive claims as compared to the modal realist. And one could most definitely argue that adopting this relatively subdued perspective is actually less useful when it comes to understanding the way in which the notions of possibility and necessity function. Although we have seen earlier that opting for a niche understanding of the latter could be equally as unbeneficial.

On this view, a possible world is just a State of Affairs that possesses the characteristics of being both possible and total. In other words, a possible world is consistent and thus formulated in such a way to where every possible and conceivable way in which that world can be is settled. Which means that not a single detail can be added to a possible world without denying it of its consistency. And just like the concretist, abstractionists believe that all possible worlds exist in just the same manner as our actual world. However, this is not to say that they consider the actual world to be a mere *example* of what form a possible world can take. On their view, the term 'actual' holds

more power, thus, the actual world serves as the baseline, the limit of which can be extended and added to, forming other possible worlds.

The existence of individuals is another important part of possible world semantics, that I shall now address. On the abstractionist conception, individuals cannot exist in possible worlds because States of Affairs are abstract entities. Which only makes sense to say that they can't exist in the literal, mereological way, as integral parts of the all-encompassing whole. The abstractionist is then compelled to define individual existence in a world as a special case of the inclusion relation, which states that, if we have two sets A and B, set A is included in set B if and only if all the elements of A are also elements of B. In other words, an individual *exists in* a possible world if the definition of that world includes that individual's existence.

Another important difference between modal realism, or concretism, and abstractionism is that individuals don't necessarily have to be worldbound. Meaning that many different worlds can exist, each of which includes the existence of one and the same individual. Furthermore, most, if not all, abstractionists commit themselves to transworld identity which is the idea that an individual existing in multiple possible worlds inherently possesses different characteristics than its counterparts and can consequently exhibit radically different properties in different worlds.

Now, we know that modal realists regard the term 'actual' as an indexical. On the other hand, abstractionists see actuality as a special *property* that serves as the distinguishing factor separating just one possible world, the actual world, from all other possible worlds. That is, the actual world represents the only way things could have been, namely, the way things currently are as a whole. Which leads us to actualism …

Actualism is the view that everything that currently composes reality, everything that exists, is actual. Utilizing the terminology of possible world semantics, everything that exists in any world, even worlds unknown to us, exists in the actual world. In contrast with this view is possibilism, which states that everything that exists in other possible worlds other than our actual world, doesn't actually exist. Instead, those non-existent objects and individuals can be referred to as *mere possibilia*. Because modal realists believe that individuals are worldbound, everything that exists in the actual world must remain distinct from that which exists in other, nonactual (merely possible) worlds. Which leads to the realization that concretism, and really basic possible world semantics in general, is committed to possibilism and the problems associated with it. Belief in abstractionism allows one to circumvent that commitment.

New actualists, who comprise a subsection of actualism, assert that individuals are and should be treated as actually in existence but not necessarily concrete. Haecceitists agree. A haecceity is a property of an individual, namely, the fact that it is that very individual possessing the property of being itself. Which leads one to say that individuals in possible worlds are not necessarily worldbound, unlike the beliefs of modal realists. This is due to the fact that as long as an individual possesses the property of being itself, it can exist in as many worlds as there are. An individual's counterparts do not inherently possess different properties, for if they did haecceity would be lost. That individual (and its counterparts) would cease to possess the property of being the same unique individual that exists in all possible worlds. And lastly, strict actualists say that if an individual failed to exist in one or another possible world, we would not recognize its absence. That is, the individual simply would not have existed, regardless of whether it is a concrete entity or an abstract conception. And it wouldn't make sense for time to track its nonexistence.

So, modal realists like Lewis believe that there exist concrete possible worlds, and our actual world is just one such world. Abstractionists, and more specifically, actualists believe that everything that exists in any possible world exists in the actual world. Which still allows us to imagine abstract worlds comprised of individuals, objects, and events that could have possibly been actualized, but were not. Let us now recollect the conception of a world that was introduced at the beginning of this paper: a realm encompassing both the physical (spatiotemporal) universe and the metaphysical universe. Which allows us to have physical instantiations of individuals as well as the thoughts that occupy their minds. My intuitions, as I'm sure was quite obvious throughout this piece, align more closely with the abstractionists, although by no means entirely. It makes sense to say that our actual world exists, and all talk of other possible worlds is just theoretical. However, I do not agree that the world that concretely exists on this picture is the physical, spatiotemporal universe. And it all comes down to truth. Some truths are undeniable, like the example of a round circle spoken of earlier. The conception of an angular circle is a straightforwardly undeniable falsity. But these analytic truths are not applicable to the objects, individuals, and events that exist in physical reality because nearly everything could have been otherwise. It is entirely conceivable that I could have been a completely different person, leading a completely different life. History did not have to necessarily play out the way it did, and the future is still up in the air, waiting to actualize itself. And even that which has already actualized itself remains a mystery which the hard sciences attempt to discover. Now, I'm not saying that the

actual world is the metaphysical universe, the realm of abstract objects and theoretical notions. But rather, a more specific version of that realm … Platonic Heaven comprised of the Ideas or Forms, if you will. Because the only way to achieve the truest and purest form of actuality, one that leaves no room for possibility, is to turn to the essences. This is where absolute truth exists even if it remains currently inaccessible to our minds. In this case we are left with a world entirely made up of the essential natures of each and every one of us and everything that occupies our environment. The entities in such a world might be abstract, but they are ideal, thus possessing a sort of truth that transcends that of analytic or logical necessity. The Ideas make everything in the realm we are exposed to the sort of thing that it is. And to a certain degree this applies to our own self, our identity. The essence of humanity is encapsulated in this world, and there it shall remain. So Lewis was right, yet so were the abstractionists. This world is a possible world similar to worlds of fiction, worlds of our creation. On the other hand, there exists only one actual world. The only issue is that it is not the one we inhabit. It is something so much bigger, a world we most likely will never be able to enter. And while it might seem as though we will always be stuck in a realm of possibility blocked from the truth, something lies in the frustration associated with our desperate attempts to represent the actual world. Even though we know things could have, and probably should have, been otherwise, here we persist. Exhausting possibilities … edging ever closer to reality.

BIBLIOGRAPHY

Lewis, David K. (1986) *On the Plurality of Worlds* (B. Blackwell)

Plantinga, Alvin. (1974) *The Nature of Necessity* (Clarendon Press)

Menzel, Christopher. (2023) "Possible Worlds," *The Stanford Encyclopedia of Philosophy* (Fall Edition)

A Look at the Paradox of Fiction in the 21st Century

Austin Straley

The paradox of fiction originates from Colin Radford's 1975 article, "How Can We be Moved by the Fate of Anna Karenina?" In the article, Radford is concerned with explaining the phenomena of why we respond emotionally to works of fiction, not what fiction is; however, it is a significant question to answer on what it means for something to be fictional if we want to give a satisfactory answer to Radford's paradox. This paper is concerned with whether there is such a satisfactory answer to the paradox of fiction that focuses on explaining the nature of fictional objects, and I will be analyzing several philosophers who have written immensely on this topic. These philosophers are Colin Radford, who posited the paradox originally; Kendall Walton, who offers a powerful theory of fiction as *make-believe*; and Craig Bourne, who offers a theory using possible-worlds semantics to explain how a fictional object, like Sherlock Holmes or Adrian Monk, can be talked about as if they were a real object.

I will first introduce the paradox in the following paragraphs, then I will reconstruct Radford's original argument in the first section of this paper, then I will look at Walton's theory of fiction in the second section, and lastly, I will argue what I believe to be the strongest solution to the paradox, which is that fictional objects are real objects on real possible worlds. The goal of this paper is to clearly articulate how to solve Radford's puzzle by offering an explanation of what fictional objects are and how they can be considered to exist. Many philosophers have offered theories of how our emotions can be rational in spite of Radford's paradox, like Martha Nussbaum and Stacie Friend, but few have successfully posited a compelling solution that doesn't require an elucidation of our emotions. Ultimately, in my conclusion, I show that such theories carry serious burdens that seem worse than the paradox itself.

The problem that Radford points out is that when we engage with a work of fiction we respond emotionally to it. Our emotional response is, without a doubt, toward the fiction and not the medium in which the story is presented. This much is obvious, as we do not scribble over the lines of a book and try to change the story, nor do we run onto the stage and interrupt a play; however, what is not obvious is what justifies an emotional response in response to the fiction.

Historically, there have been two traditions for answering the paradox: the normative and descriptive responses. The paradox can be put into three premises:

> P1. We experience (genuine, ordinary) emotions toward fictional characters, situations, and events.
> P2. We do not experience (genuine, ordinary) emotions when we do not believe in the existence of the objects of emotion.
> P3. We do not believe in the existence of fictional characters, situations, and events. (Friend 2016, pp. 1-2)[1]

The difference between the normative and the descriptive responses is whether one wants to accept all three premises or reject one or several of the premises. In the following two paragraphs, I will briefly provide a history of the two traditions and where the current debate stands.

The tradition sired by Colin Radford is called the normative tradition, which accepts all three premises as true and concludes based on its implication. Radford is concerned with whether our emotional responses to fiction can be considered rational, not the ontological status of fictional objects or the metaphysics of what it means for something to be fictional. I will review Radford's argument in more detail in the next section.

The descriptive response is one in which a premise is rejected. By far, the most rejected premise is P2; this is, on its own, the most dubious-sounding premise. However, there is a good reason not to reject it—we do not need to dig ourselves out of the metaphysical hole of fictional object ontology to talk about our emotional response to fiction. The paradox arises regardless of whether P2 posits the existence or inexistence of fictional objects—we do not need to believe in fiction to feel an emotional response. Philosophers concerned with the descriptive response are more concerned with the ontology of fiction than with dissolving the paradox, which can be seen clearly with the philosopher Kendall Walton, who contributed immensely to the philosophy of fiction.

Walton rejects the first premise, arguing that we do not experience genuine emotions towards fictional things; rather, we experience quasi-emotions (Friend 2016, p. 1). Walton offered a theory aptly named the make-believe theory: When we engage with fiction, we are really playing

a game where we make-believe that certain things are true. In a game of make-believe, our emotions are not fake or pretended; rather, they are real in the context of the game but not real or genuine outside of the game. Walton's theory will be revisited in the second section of this paper in more depth.

SECTION 1: RADFORD'S ARGUMENTS AND WHY HE BELIEVED EMOTIONS ARE IRRATIONAL

After proposing and refuting five other possible solutions, Radford's conclusion was to accept that our emotional responses to works of fiction are simply irrational—that there is no explanation possible. Radford makes several arguments that challenge what I consider to be the intuitive notions of resolving the paradox. I will briefly introduce these arguments, showing how Radford reached his conclusion.

Radford first proposed that our emotional response to fiction is because we are caught up in the story and forget that it is not real. Radford jests that we do not jump on stage, nor do we scribble out the text and rewrite the book. We may feel like we suspend our disbelief and are engrossed in a fictional work, but that does not mean we forget it is a story. If one were to receive a phone call while in the audience at a play, it would not be like they would suddenly remember they were not a part of the fiction. We are always aware that we are engaged with some medium that displays fiction, but that does not take away from our ability to experience the fiction independently of that medium.

His second suggestion is to ignore our doubts about the story's reality willingly. Building on the previous idea, it is clear we do not believe the events in a play are actual. Instead, we immerse ourselves in the story by setting aside our skepticism. This concept is similar to his first suggestion but with a twist: rather than fully believing in the fiction, we engage with it by holding our disbelief at bay. Radford initially talks about plays, but movies provide a more relatable example. Consider watching a horror film like *Alien*. Do we forget that we cannot just skip the frightening scenes? As someone who gets easily scared by horror films, I always keep the remote control close. However, the mere presence of the remote suggests that we do not fully immerse ourselves by suspending disbelief, indicating that this is not the sole reason we react emotionally to Ridley Scott's iconic film. Some might argue that this is not a fair assessment of suspending

disbelief; nevertheless, it is impossible to completely set aside our disbelief, meaning there must be a more compelling reason for our emotional responses to fictional stories.

The third solution that Radford entertains is that "it is just another brute fact about human beings that they can be moved by stories about fictional characters and events" (Radford 1975, p. 72). Radford rejects this by claiming that not everyone reacts emotionally to works of fiction, which he implies is a stronger question than why we react emotionally to works of fiction. This argument that he rejects is what I consider Radford's most dubious argument here, as he would rather settle on rejecting the rationality of our emotions than accepting that it is just a brute fact of being human that stories move us. While this is not a perfect argument or solution, it carries significantly less baggage than his conclusion that emotions are irrational. Radford wants to avoid this solution simply because he thought it would be harder to explain why there are idiosyncratic responses to fiction—especially when there is no emotional response by the person engaged with the fiction.

The fourth solution Radford entertains is that an availability heuristic occurs when we engage with fiction, e.g., the story has some violent psychopathic serial killer that stalks people who leave their windows open. In that case, the person who engages in the fiction might think that there is an increased likelihood of being a murder victim if they did something similar as the fictional character. Radford posits that we might feel an increased likelihood of events happening considering engaging with fiction, but this is demonstrably false. For example, we do not think that the problems that occur in *Hogwarts* are likely to occur to us because they are too fantastical.

Lastly, the fifth solution that Radford gives, the last one before he resigns the rationality of emotion, is that when we "weep for Anna Karenina, we weep for the pain and anguish that a real person might suffer and which real persons have suffered, and if her situation were not of that sort we should not be moved" (Radford 1975, pp. 74-75). Radford argues that we do not react to the circumstances of fiction and feel a particular emotion because we imagine those circumstances affecting a real person; rather, the object of our emotions is the fictional character.

Radford posits that we have an emotion toward a fictional character because we imagine that if it were someone else who is real and not imaginary, then we would be justified in having an emotional reaction because they are real. This argument is obviously dubious, something that Radford likely knows that one should not agree with in the first place because this solution is simply question-begging that we would be justified to have an emotion towards a real person in

place of a fictional character if such (real) person were going through the same plight as the fictional character. The absurdity here is that if we took Ellen Ripley, the brave hero of the *Alien* movie, and imagined the plight that she faced in the fiction as something that could happen to a real person, then we would not be justified in believing the emotional responses to the awe-inspiring acts of Ellen Ripley if we think it had happened to Sigourney Weaver. If this were the case, then it would be clear that whatever is eliciting the emotional response in the fiction is not the characters themselves.

Radford notices something in his fifth solution that seems to explain the paradox but effectively resigns to his sixth solution, saying, "Perhaps there really is no problem…I am left with the conclusion that our being moved in certain ways by works of art, though very 'natural' to us and in that way only too intelligible, involves us in inconsistency and so incoherence" (Radford 1975, pp. 75-78). The complexity of our emotional response to fiction is why Radford is surrendering the rationality of emotions. As Radford shows throughout his solutions, we can give an effective response to why a particular fiction moves us, but we cannot explain how fiction can move us. How does the character Anna Karenina give rise to an emotional response? Explaining the complexity gives rise to a harder question than Radford initially asked: What does it mean for something to be a fictional object? In answering this question, one can hope to find an answer to the paradox that Radford gives. In the next section, I will go over what seems like the strongest theory of what it means for something to be fiction.

SECTION 2: FICTIONAL ANTI-REALISM VS. FICTIONAL REALISM

There are two theories of what fiction is. The first is Fictional Anti-Realism, which holds that fictional objects do not literally exist. The second is Fictional Realism, which holds that fictional objects exist in some way beyond our imagination. I have already introduced Kendall Walton, a notable proponent of the former view. I will briefly surmise Walton's theory and show why it is a good theory of fiction. I will argue that theories like Walton's, even if they are likely involved in our understanding of how we interact with fiction, are inadequate in articulating a solution to Radford's paradox. Then, I will introduce a promising theory of fictional realism that involves the notion of possible worlds, often attributed to the philosophers David Lewis and Graham Priest. This theory is Craig Bournes', who is not concerned with the paradox of fiction but rather a way

of explaining time in fiction; however, his argument of what a fictional object is will solve Radford's problem at a high cost.

In the introductory section, I discussed the two traditions of responding to the paradox using Walton's and Radford's approaches. However, what is important to discuss is the nature of fiction itself. Many philosophers of literature and fiction have offered in-depth explanations of how our emotions are rational regarding fiction, but these arguments do not give us a satisfactory account of what fiction is or what a fictional object like Sherlock Holmes is.

Returning to Kendall Walton, whom I discussed in the introduction, what makes something fictional is whether one can engage in a game of make-believe with it (Walton, 1990). On Walton's account, the norms of responding to fiction play a significant role in what constitutes a fictional story. The constitution of fiction is found in the participatory nature of playing a game rather than in the story itself. Further, Walton believes that the emotions that we have towards fictional characters, in the game of make-believe, are real, genuine emotions, but they are of a diminished nature—the fear that occurs in watching a horror film is quasi-fear, which is fear-like.

However, there are problems with Walton's theory. While I consider it to be a strong theory in explaining how we interact with fiction, it does not provide an intuitive notion of how we can say statements like 'Sherlock Holmes is a fictional character' without some significant paraphrasing. When we say something like 'Sherlock Holmes is a fictional character' what we really mean is that when one engages in a game of make-believe, there is a name that refers to a particular prop in that particular game of make-believe. This is a significant amount of maneuvering to explain something that seems like an ordinary, simple proposition. This can be further complicated by comparing fictional characters. For example, if someone were to say, 'Dr. House is a lot like Sherlock Holmes' or 'Dr. House is based on Sherlock Holmes.' In ordinary discourse, it seems more intuitive to consider these fictional characters as independent objects rather than names of make-believe props. This may not be a significant problem for Walton and fictional antirealists, but it seems like we lose the ability to refer to what seems like distinctive characters in stories.

SECTION 3: FICTIONAL REALISM AND POSSIBLE WORLDS

Fictional antirealists like Walton offer powerful theories for explaining how we interact with fiction, and fictional realists like Bourne agree that Walton's theory of make-believe is most likely

partially correct (Bourne 2016, p. 13). However, there are other explanations for what constitutes a fictional character. Again, take the proposition 'Sherlock Holmes is a fictional character.' This proposition implies that some object exists and bears the property of being fictional, which brings us back to needing a way to explain fictional ontology—if there were a way, we could describe a fictional object as being like a real object, then there would be no paradox. Essentially, if one could be a fictional realist about fictional characters—meaning that fictional characters are an object of some kind—then we can reject the third premise of Radford's paradox from 'we do not believe in the existence of fictional characters, situations, and events' to 'we believe that fictional characters exist as abstract objects.' If fictional characters are abstract objects, then they are non-physical, non-mental objects that exist outside of space and time. If we think of fictional characters as abstract objects, then there is no paradox because it is not irrational to have an emotion towards an existing object. For now, we should ignore how an abstract object can cause an emotional response in a person.

How do these supposed abstract objects come about? There are two theories for how. One is that an author, or fiction-maker, creates the fiction, thus creating the abstract object of a particular fictional character. Another is that all fictional character exists already, not requiring an author to author them into existence, and that there are possible worlds in which each fictional character exists and has particular properties. Philosophers who endorsed the former are Amie Thomasson and John Searle, and philosophers who endorsed versions of the latter in some ways are David Lewis, Graham Priest, and Craig Bourne.[2]

Briefly, the former view carries a significant flaw. What if an author creates every fictional story? If an author authoring is all it takes to bring an abstract object into existence, then there is a problem with overlap. Two short stories by Jorge Luis Borges are helpful here. They are *The Library of Babel* and *Pierre Menard, Author of the Quixote*. Both stories are centered around fictional works and the notion that authorship is relevant to the fiction. In *The Library of Babel*, there exists an infinitely sized library that contains every concatenation of letters and words, thus, one must imagine that every possible book that could be written exists within the library.[3] That means there can be no new fiction or stories since everything is already contained in the library. While it's true Borges' story is simply a fictional story, there is a real-world comparison to be made. A fan of Borges made a digital library, that is completed up to 5000 words, where any string of words and letters can be found on a possible "shelf."[4]

The alternative to the view that we create fictions is known as possibilism.[5] This is where I believe the strongest solution to the paradox lies. Possibilism holds that fictions are just iterations of a possible world—the way the world could have been. For example, there could have been a Sherlock who was a detective, or there could have been a Dr. House who was infamously cruel to his patients. Bourne endorses a stronger view of possibilism, holding that fictional worlds are a subclass of possible worlds and that fictional characters are occupants of these worlds. Moreover, Bourne endorses a modal realist view of these possible worlds, meaning that these worlds are real and concrete and inaccessible to us—a view like that of David Lewis (Bourne 2016, p. 16).

Where Bourne differs from Lewis is on the nature of fiction. Lewis subscribes to a fictional antirealist view, even though he'd likely accept that many fictional stories happened in some modal world; rather, he would argue that fiction-makers engage in pretending to report facts about how things are (Bourne 2016, p. 21). Bourne also wants to reject Lewis's theory of fiction because Lewis does not offer a satisfactory account of impossible fictions, because for Lewis, impossible things cannot occur in a possible world. Bourne instead holds that what can be true in a fiction can be true only in a single possible world, making it the case that there can be no impossible instances even for fiction. To elaborate on this, suppose there are two similar fictional stories, each one with their fictional world. Let one be the fictional story of *Sherlock Holmes*, the 2008 film with Robert Downey Jr., and the other be *Sherlock*, the TV series that aired from 2010 to 2017. On Bourne's account, there is a fictional character in both fictional worlds—and fictional worlds are possible worlds—what can be true of the character Sherlock Holmes is determined by the possible world itself.

Bourne also distinguishes his view from that of Graham Priest's view on fictional worlds. Priest posits that fictional objects are non-existent objects and that we can have intentional states in which the mind is directed at these non-existent objects (Priest 2005; Bourne 2016, p. 22). Moreover, Bourne's view differs from Priest's, specifically in the notion of whether possible worlds exist. For Priest, non-actual worlds do not exist; rather, they are non-existent objects, as opposed to Bourne, who holds that possible worlds, which includes fictional worlds, exist as concrete objects.

Accepting a notion of modal realism seems like a heavy burden, though. It doesn't seem intuitive to say that the events in the movie *Alien* occurred in some world. This is a similar type of baggage that Walton's theory had, where we must accept that when we say something like

‘Sherlock Holmes is a detective’ we really mean something like, ‘in a particular game of make-believe when one says the name ‘Sherlock Holmes’ one really means to refer to a prop in a game of make-believe.’ Both theories resolve the paradox but lack parsimony in explaining our faculty for engaging with fiction, and how we can construct a fictional work.

Ultimately, there appears to be solutions to the paradox of fiction, yet the two strongest theories do not seem satisfactory to me. Walton’s theory imposes a burden of paraphrasing that we don’t mean what we say in regard to fiction, and Bourne’s theory leads us to accept that we really mean what we say—so much so that we believe in the fiction’s occurrence as a fact. Whether the paradox can be solved adequately without an appeal to the nature of our emotions is up for debate still, but as shown in this essay, there is a problem of accepting a heavy ontological burden in articulating a view in which fictional objects can exist.

NOTES

1. Friend is following a tradition of paraphrasing Radford’s initial argument, but it is important to be clear that Radford is not very explicit. Many philosophers have argued against Radford’s puzzle, often listing his argument similarly to how Friend detailed the premises. This should not be an issue of concern, as when one talks about the paradox of fiction, they often talk about these three particular premises.
2. Importantly, these three philosophers differ on what exactly is a possible world and whether these possible worlds are abstract objects. For Lewis, he believes that possible worlds are real, spatiotemporal worlds that exist. For Priest, he disagrees that these possible worlds exist spatiotemporally.
3. This “library” actually exists online now, but only up to 5000 words. Any and all fictions that are less than 5000 words, regardless if they’ve been written yet, can be found on the library which gives an algorithm for what “shelf” that particular story would be found.
4. See https://libraryofbabel.info
5. See https://plato.stanford.edu/entries/fictional-entities/

BIBLIOGRAPHY

Borges, Jorge Luis. (1998) *Collected Fictions.* Trans. Andrew Hurley (New York: Viking)

Bourne, Craig, and Emily Caddick Bourne. (2016) *Time in Fiction* (Oxford: University Press)

Friend, Stacie. (2016) “Fiction and emotion,” in: Amy Kind (Ed), *The Routledge Handbook of the Philosophy of Imagination*, pp. 217-229 (New York: Routledge)

Radford, Colin, and Michael Weston. (1975) "How Can We Be Moved by the Fate of Anna Karenina?" *Supplementary volume - Aristotelian Society 49.1*, pp. 67–93

Walton, Kendall L. (1990) *Mimesis as Make-Believe: On the Foundations of the Representational Arts* (Cambridge: Harvard University Press)

GOODNESS WITHOUT GOD

Arwa Hammad

INTRODUCTION

I am certain that anyone reading this paper has experienced the sting of a harsh insult or the warmth of a sincere compliment. These statements were either morally "good" or morally "bad." But do these moral values stem from objective rules, or are they subjective constructs influenced by individual beliefs and cultural norms? If you were to listen to the Christian apologetic (henceforth C.A.), they would tell you that they are rooted in objectivity and are facts we follow that are ordained from God. Otherwise, there would be no way we would have an objective moral reality (henceforth OMR). While I agree that there is an OMR, I disagree that it is divinely mandated, and claiming it as such proves problematic for such a reality. In this paper, I aim to prove that this OMR, first, exists, and second, stands independently of theological roots while still transcending individual beliefs and cultural divergences. I will do this by explaining the general C.A. argument using Richard Swinburne's *The Existence of God* book to truly understand the perspective of those who believe in the divinely mandated OMR and their arguments on why the atheistic perspective is not the best explanation for our OMR. Finally, I will explain the counterarguments I expect to arise and comment on them to solidify my argument. With this said, I will be dividing my paper into five sections. Section I will cover common terminology and explain each term so there is a basic understanding. Section II will cover the C.A. argument, and Section III will cover specific issues regarding the C.A. argument. Section IV will establish my definition and criterion for an OMR without God, and Section V will cover possible counterarguments stated by the C.A. and my rebuttal to them.

SECTION I: WHAT'S MEANT BY…

To truly understand the arguments brought by the C.A., we must understand the terms commonly used and offer standard definitions so arguments can be understood and moved forward. The first term to define is *morality*. When I say morality, I mean codes of conduct followed by societies, groups, or individuals that guide us to do what is innately understood as good and bad. For example, picking up a piece of trash from the floor would be good, and stealing from the poor for

fun would be bad. For morality to be objective, there must be two conditions followed: that moral rules must apply regardless of our endorsements and regardless of whether they get us what we want, and that they normatively apply to every person insofar that, when presented with certain specified situations, all rational persons should follow the same moral guidelines and act relatively the same.

Finally, the last term that needs to be explained is God. I will be adopting Swinburne's definition of God, which states that God is:

1. Eternal, in that "he always has existed and always will exist" (Swinburne (1979), p. 7). This should not be confused with being "timeless" or "outside time," as it is sometimes interpreted within Christian tradition.
2. Omnipresent, in that he is a spirit that has always existed and always will exist.
3. Perfectly free, in that God's actions are not causally influenced by any "object or event or state (including past states of himself)…" Rather, God's choices at "the moment of action alone determines what he does" (ibid., p. 7).
4. Omnipotent, in that God has maximal power and can do anything, even create logically incoherent things (such as a round square).
5. Omniscient, in that God has maximal knowledge, essentially knowing "whatever it is logically possible that he know" (ibid., p. 7).
6. Perfectly good and has *perfect goodness*, in that God always does what is "morally best" (when such actions exist) and does not engage in any "morally bad action[s]" (ibid., p. 7). This is where moral realism comes in.
7. The creator of all things, in that everything that exists (apart from God) exists because God either makes it exist or permits it to exist.

For the argument's sake, Swinburne posits God's definition as a "brute fact" that simply does not have an explanation. Therefore, when Swinburne refers to God, these values are stipulated.

SECTION II: WHAT THE C.A. POSITS

When the C.A. argues that morality necessitates God, they assert that the presence of objective moral standards and responsibilities depends on God's existence. Without God, objective moral

values and obligations would not exist. Therefore, the C.A. believes morality is not a social construct or a result of human development but a necessary behavioral standard God gives us that surpasses cultural and individual variations among people. These moral principles are inherent to reality and not dependent on human opinion or cultural conventions. Finally, all moral actions are either inherently bad or good to do. For example, specific crimes, such as murder or theft, are universally immoral, regardless of individual beliefs or cultural practices.

Before discussing the argument for the OMR, it is crucial to outline Swinburne's argument for the presence of God, which is based on moral realism, probability theory, and inductive reasoning. Swinburne distinguishes between two types of *inductive arguments*, or arguments that use specific observations or evidence to make generalizations or predictions about broader situations or patterns: P-inductive and C-inductive. In a correct C-inductive argument, the premises "confirm" the conclusion, making it "more likely or probable than it would otherwise be" (ibid., p. 6). In a correct P-inductive argument, the premises "make the conclusion probable" in that the probability that the conclusion is true is greater than 0.5. Swinburne stresses the importance of ensuring that the premises of an argument are known to be true by those disputing the conclusion. He illustrates this with an example related to the existence of God, stating:

P_1: If life is meaningful, God exists.
P_2: Life is meaningful.

C: God exists.

This is a valid argument, according to Swinburne, but is essentially ineffective, as both sides of the debate do not commonly accept the premises in terms of the existence of God. Swinburne states that "[w]hat are clearly of interest to people in an age of religious skepticism are arguments to the existence (or non-existence) of God in which the premises are known to be true by people of all theistic or atheistic persuasions" (ibid., p. 6). So, for an argument for God to be effective, not only should its premises "probabilify" the conclusion, but they should also be known to be true by those disputing it. Therefore, to prove the existence of God, Swinburne will solely focus on *a posteriori* arguments, which are based on experience and facts, and not on *a priori* arguments, or arguments derived from logically required truths (ibid., p. 8). By focusing on a posteriori arguments, he aims

to ground his analysis in tangible human experiences that can be observed or inferred from the world around us.

With that understanding, Swinburne creates his probabilistic argument by utilizing the Bayes Theorem, which describes the probability of an event based on prior knowledge of conditions that might be related to the event. He uses the Bayes Theorem to update the probabilities of God's existence by incorporating new evidence, and he compares the likelihood of observing certain phenomena given God's existence versus the likelihood that God does not exist. This evidence is evaluated through both P and C-inductive reasoning frameworks, and he uses these arguments to assess whether the evidence supports the hypothesis that God exists (henceforth the theism hypothesis) more than its negation. By stacking these arguments (for example, let 1, 2, 3, 4 be arguments) so that 1 supports 2, 1 and 2 support 3, 1, 2, and 3 support 4, and so on, Swinburne's trying to build up a series of C-inductive arguments that, on their own, do not witness to the likelihood of anything, but, when combined, give a sound P-inductive argument, allowing him to reach his goal of establishing the theism hypothesis (and as such, an objective OMR) is a likely conclusion (ibid., pp. 13-14). P-reasoning, the probability of his hypothesis given the evidence is compared to a threshold of 0.5. If the probability exceeds this threshold, the argument is considered strong, and his hypothesis holds.

In proving this, Swinburne introduces the *Principle of Credulity*, which states that "…if it seems (epistemically) to a subject that x is present (and has some characteristic), then probably x is present (and has that characteristic); what one seems to perceive is probably so" (ibid., p. 303). In other words, if something appears to be a certain way, it is reasonable to believe it to be so unless there are specific reasons to doubt it, thereby using Occam's Razor, which states that the simplest explanation is preferable to a more complex one. Swinburne's principle applies not only to present experiences but also to memories of past experiences. Essentially, how things seem to be or have been experienced or remembered provides a valid basis for beliefs about reality, so "…memory is to be trusted" (ibid., p. 303). Moreover, the principle emphasizes that the evidence provided by experiences applies only to what appears to be present, not to what appears to be absent. For example, if someone perceives a table in a room, it is reasonable to believe there is a table. However, if someone perceives the absence of a table, it would be insufficient to conclude there is not one unless there are additional reasons to believe so, such as thoroughly searching the room. So, if someone genuinely experiences something, such as a vision of heaven or God, it

provides *prima facie* reason to believe in the reality of that experience, and it should be taken seriously as evidence for the existence of God pending further examination or justification. Such claims can be doubted or proven false if:

1. Perceptual claims (of the relevant kind) "are generally false" in that they have a history of unreliability or if the circumstances under which the claim was made are known to produce unreliable perceptions.
2. The type of perception claimed has been shown in the past to be false or unreliable under similar circumstances.
3. Certain doubtful experiences are necessary prerequisites for making what is claimed to be the accurate, perceptual claim.
4. The claim can also be challenged by demonstrating that it is highly probable that the perceived object was not present or that something else was the cause of the perception.

Swinburne notes that some may argue that religious experience can be doubted by claiming one of these four things based on the claims above:

1. Religious perceptual claims are generally unreliable.
2. The unreliability of religious experiences is purportedly evidenced by the conflicts among them, in that "[t]he issue arises with the claims of religious experience as with the claims of miracles, whether there is a general proof of their unreliability in the fact that so many of them conflict with each other" (ibid., p. 316).
3. It is highly improbable that the claimed object of religious experience (in this case, seeing God) was present, "and so the subject could not have perceived him" (ibid., p. 319).
4. The religious experience probably had a cause other than its purported object.

To debunk each claim, Swinburne states:

1. "Most religious experiences are had by people who normally make reliable perceptual claims, and have not recently taken drugs," in that most religious experiences are had

by people who typically make reliable perceptual claims, so the truth value of their claims can be trusted (ibid., p. 315).

2. Differences in religious vocabulary and interpretations do not necessarily invalidate these experiences. People from different cultural and religious backgrounds may interpret their experiences using their own familiar religious vocabulary. For example, experiences described as encounters with different deities may be different perspectives on the same divine reality, known by different names in different cultures. He also adds that conflicts in descriptions of religious experiences do not automatically invalidate the experiences themselves. He compares this situation to conflicts in perceptual claims made by different groups, such as Babylonian and Greek astronomers describing celestial phenomena differently. Just as conflicting astronomical observations do not negate the existence of objects in the sky, "the perceptual claims of each group constituted arguments against the perceptual claims of the other group; and, given that the perceptual claims of both groups were equally weighty in number and conviction, that further arguments were needed to adjudicate between them" (ibid., p. 317). So, conflicting descriptions of religious experiences merely challenge specific detailed claims rather than casting doubt on all claims of religious experience.
3. If someone claims to have had an experience of God, the burden falls on the skeptic to demonstrate that it is highly improbable that God was actually present for perception. He argues that since God, if He exists, is omnipresent, the claim that God was not present can only be substantiated if it can be proven that it's "very probably there is no God" (ibid., p. 19). Additionally, a lack of similar experiences with God does not mean God was not present. Swinburne argues that the omnipotence of God allows for selective revelation, meaning that God could choose to reveal Himself to some individuals while remaining hidden from others. Therefore, the absence of a universal perception of God does not inherently discredit individual claims of experiencing God's presence. However, Swinburne acknowledges that the lack of widespread perceptual experiences of God diminishes the evidential weight of any individual claim.
4. While the cause of a religious experience may be due to other factors, an apparent experience of an entity is considered as such if the entity is part of the causal chain that leads to the experience. Since God, if He exists, is omnipresent and sustains all causal

processes, any experience that seems to be of God would indeed be genuine. Experiences of God can occur through two means: either God intervenes in the operation of natural laws, producing an event beyond what natural laws would typically produce, or He sustains the normal operation of natural laws that leads to religious experiences under specific conditions, such as fasting. With that, Swinburne emphasizes that the fact that natural processes brought about a religious experience of God does not discredit its validity. To disprove the veracity of such an experience, one must demonstrate that God did not cause these processes, which, according to Swinburne, can only be achieved by proving that God does not exist. In essence, religious experiences apparently of God should be taken as veridical unless it can be demonstrated, based on other grounds, that it is significantly more probable than not that God does not exist.

SECTION III: WHAT'S WRONG WITH THE C.A. ARGUMENT?

Now, with the main argument set by the C.A., what is wrong with the argument? There are four main problems:

1. To combine and respond to Swinburne's 1st and 3rd rebuttal, the Principle of Credulity is unsubstantiated; religious experiences can be influenced by cultural conditioning, confirmation bias, and/or subjective interpretation, which casts doubt on their objective validity as evidence for the existence of supernatural entities. Additionally, the argument from silence, the problem of selective revelation, the plausibility of alternative hypotheses, and the reliability of perception collectively challenge the validity of claims about experiencing God's presence.
2. The lack of empirical verification and interpretive flexibility, combined with potential confirmation bias and conflicting claims of exclusivism versus pluralism, pose significant challenges to establishing the validity and universality of religious experiences.
3. Establishing a clear causal connection between God's presence and the perception of God raises epistemological concerns about the burden of proof and the reliability of

subjective experiences alongside alternative interpretations rooted in cultural, social, or personal factors.

4. Morality reliant on God is either arbitrary or irrespective of God.

Regarding my first rebuttal, Swinburne is correct in claiming that most people who claim to have encountered religious manifestations or communications from God are trustworthy and in a condition of mental clarity, but this does not imply that what they are experiencing is what they are stating. Furthermore, he fails to admit that there is additional evidence that these divine instructions are not supernatural but rather the result of a combination of *confirmation bias*, or the tendency for people to selectively seek, interpret, and remember information that confirms their preexisting beliefs or hypotheses while ignoring or discounting conflicting evidence, and cultural conditioning. Empirical research has shown that our brains are susceptible to confirmation bias and believing what we have been taught and told. In essence, we tend to believe what we are told while simultaneously resisting changing our minds about previous findings. This encourages us to seek explanations for our experiences that correspond to the in-group's overall belief system rather than determining the true answer to what we have encountered. We are prone to *hyperactive agency detection*, or the human tendency to perceive agency or intentionality in situations where none exists, often attributing events to hidden agents or entities even without evidence. As such, we are apt to leap to conclusions and identify patterns, but we are not good at statistical thinking or evaluating what coincidence is and is not; we see patterns and establish linkages even when there is no genuine pattern.

However, even if there was no indication that the skeptic opposed the C.A. and their religious experiences, Swinburne is incorrect in claiming that the burden of proof is on the skeptic. The C.A. is also unaware of the rationale or cause of such an event, and because such proof is not available, all they know is that they had an experience. The C.A. must establish that their alleged experience with God was caused by God sending a message. It is not up to the skeptic to disprove the claim. To illustrate this, imagine that person B claims to have seen a UFO in the sky and believes it was an encounter with alien life. Person B insists that their encounter demonstrates extraterrestrial presence. While I do not doubt that person B believes they have seen a UFO, this does not imply that they did. To show this, person B must present proof to back up their claim that the UFO was of extraterrestrial origin. Skeptics are not responsible for disproving person B's

statement since the burden of proof is with person B to present convincing evidence or explanation for their belief in extraterrestrial visitation. Without such proof, the assertion is baseless and subject to question. In my second rebuttal, Swinburne's assertion that differences in religious experiences do not undermine the existence of God is problematic. As previously argued, while individuals may honestly report their religious experiences, the interpretation of these experiences is inherently subjective and cannot be definitively verified. For instance, if person A dreams of encountering the Christian God affirming Christianity as the true religion, while person B dreams of encountering Allah affirming Islam, how can both interpretations be reconciled as equally valid? Such discrepancies raise questions about divine revelation's fairness and subjective experiences' reliability. Accepting both interpretations as equally valid implies divine misguidance or the dismissal of anecdotal evidence, neither of which are satisfactory positions for the C.A. to hold. Swinburne now must claim that we cannot trust anecdotal evidence, even from trustworthy people, or that God is purposefully swaying certain individuals the opposite way with clear messages, such as with person A and person B.

My third rebuttal is connected with the other rebuttals, as Swinburne claims that if someone experiences what they believe to be God, it is considered genuine if God is somehow involved in causing that experience, as God is everywhere and controls everything. Therefore, he could create experiences of Himself through natural processes or by directly intervening in the world. However, as mentioned before, religious experiences can either be trusted, and God seemingly deceives certain people, or they cannot be trusted at all. Even if we contended that trustworthy people are having these experiences, they can be adequately explained by natural phenomena like psychological mechanisms, cognitive biases, or altered states of consciousness. Without robust evidence to differentiate between supernatural and natural explanations, simply attributing religious experiences to God is not valid. Additionally, Swinburne's assertion assumes a clear and direct causal relationship between God's presence and the perception of God, which presents an ambiguity in causal chains. Establishing such a connection is challenging as it requires reliable methods for identifying and verifying the involvement of supernatural entities in subjective experiences, which can be subject to interpretation and bias.

With the first three rebuttals, it is clear that Swinburne does not adequately prove his theism hypothesis. Morality inherently consists of qualitative standards, making it challenging to calculate a quantitative outcome using a precise formula where factors can only be assessed subjectively.

He aimed to demonstrate that a certain probability exceeds 0.5, but his hypothesis does not inherently involve any quantitative elements for him to prove it. In other words, his utilization of Bayes' Theorem is not enough to create a probable cause for God or an OMR, as the impression of accuracy is comparable to utilizing a scale to determine the length of an object; he needs to find an OMR using the appropriate tool, which cannot be found under scientific or empirical lenses. Furthermore, his capacity to compute the probability of God's existence does not imply that God exists. While he states that he aims for a "probable conclusion, not an indubitable one" (ibid., p. 2), I argue that his argument for God does not prove probable. Notwithstanding the problems with his premises, he oversimplifies his entire argument.

He consistently employs this idea to reject alternative theories as more intricate and less likely. Defining simplicity is subjective and may be projected onto the environment rather than discovered within it, making it fundamentally subjective. For example, a physicist perceives quantum physics as simpler than an elementary student would, showing the varying definitions of simplicity. Swinburne does not have an exact measure of simplicity that most can agree with, which makes his argument foundationally weak. He exchanges a hypothesis with a high prior probability but low likely (random chance) for a hypothesis with a high likelihood but low prior probability (a conscious creator). Given that random chance is a basic process, while a conscious creator is significantly more complex, and that the likelihood of all explanations decreases exponentially with complexity, it is evident that the probability of a conscious creator is not just lower but substantially lower. A finite range might be deemed "small" compared to real numbers' limitless reach. There is no absolute measure of largeness without a reference point to known potential values, making the observation worthless.

His argument is a *fine-tuning argument* due to his oversimplification, where the cosmos' basic constants and beginning circumstances are precisely adjusted to allow for the creation of life. He contends that the specific values of basic constants and beginning circumstances required for life are unlikely to occur naturally but are more likely if God exists. The issue here is that any argument claiming the fine structure constant falls within a "small range" could also suggest what other values it could have been, not just what it is. This makes his argument untenable. Swinburne should have anticipated an exceptionally low likelihood of any specific event occurring because of the vast range of possibilities in the universe, and his argument suggests there were limited choices to start with.

While on the topic of oversimplification, Swinburne contends that the concept of *infinity* is simple since he suggests that the characteristics of God and his perfections may be seen as infinite values along those dimensions. He compares these infinite "values" to numerical values, arguing that whereas a discrete number always prompts questions about its precise limit, the idea of infinity— boundlessness—avoids such inquiries (ibid., p. 97). This viewpoint suggests that people's experiences, closely linked to specific numerical values like speed, do not adequately equip them to understand the abstract concept of infinity. Swinburne continues to assert his thesis, promoting the idea of the simplicity of infinity in contrast to commonplace quantitative norms.

Swinburne's claim that God can bring about any logically feasible situation leads us to a dilemma. If we accept the idea that each of God's abilities is initially different and autonomous, we are forced into a paradox. Comparing the abundance of divine talents shows a lack of adherence to Swinburne's simplicity requirements since the limitless collection of these supreme traits contradicts the idea of simplicity he supports. Infinity, along with the idea of limitless divine characteristics, raises skepticism about the simplicity proposed by Swinburne. This idea reveals a more complex world than the simple infinity implied, similar to the vast array of abilities associated with a deity. Swinburne's recurring claims about the simplicity of his hypothesis based on infinity confront a significant problem, as shown above. Upon deeper examination, the proposed connection between the limitless essence of the divine and conceptual simplicity shows notable flaws; the core components of his argument are weakened due to the intricate complexity of infinite qualities being revealed.

Additionally, his explanation of God's infinite capabilities exemplifies simplicity. While the idea of a being having an unlimited array of traits, each operating within its area but collectively contributing to a unified nature, does not weaken the notion of divine simplicity; it arguably emphasizes a different interpretation of simplicity beyond our limitations and knowledge. When seeing God's characteristics as eternally diverse and numerous, comprehending these traits within the framework of 'simplicity' becomes more intricate. As such, some questions arise: Does simplicity indicate an absence of internal divisions or distinctions when referring to the divine? Or does it imply a more advanced level of organization that can handle complexity inside a single system? The problem is understanding an endless collection of unique elements, which Swinburne has failed to consider.

Even if we ignore the problem of simplicity in proving the existence of God and an OMR, having a belief in God proves problematic for an OMR. In addressing the fifth and final point, I would like to first examine certain moral maxims that we, as a society, virtually all agree with:

1. It is wrong to enslave others.
2. It is wrong to kill innocent children.
3. It is wrong to steal from people.

The issue with the C.A. argument is that since morality depends on God, it is subjective to God. There is no objective way to assess God as good or bad; it is what it is, and we then define morality around Him. This presents an issue, as Plato's *Euthyphro* demonstrates: Is something commanded by God because it is good, or is something good because God commands it?

If the C.A. accepts the first reason, morality is separate from God, introducing the *Divine Command Theory* (henceforth D.C.T.). With this theory, the C.A. must accept that morality is something out of God's control, which goes against the very definition of God. Therefore, the C.A. must accept the second option. However, if whatever God commands is good, then morality must be arbitrary. God could have very well commanded that torturing babies is morally good, and it would be this way independent of any external version of morality. To say that morality depends on the existence of God is to say that none of these moral judgments is true unless God exists. With that, if God turned out not to exist, slavery would be okay; murdering innocent children would be okay; stealing would be okay. In essence, the pain of another human being would mean nothing if God never existed.

Finally, the C.A. must acknowledge that our moral value is contingent on the presence of God, implying that individuals are inherently valueless and the care we have for their welfare holds no greater ethical importance than the concern some have for material possessions. If we accept the C.A. perspective, we would need to acknowledge that our care for other people is justified solely because they are valued by someone else, in this instance, God. The C.A. argument suggests that only God's importance is fundamental, rather than considering another being's anguish or the wellbeing of a loved one as reasons for providing help or making sacrifices. The objectivity of the C.A.'s morality is irrelevant as there is no evidence to support their moral nature.

SECTION IV: SO, WHAT'S THE OMR?

Having established the issues latent in the C.A. argument, I can explain how an OMR comes to exist without God. First, to establish an OMR, I must establish:

1. What moral facts are.
2. Moral facts exist independently of beliefs, biases, and opinions.
3. The means of achieving such moral facts are independent of beliefs, biases, and opinions.

In addressing the first point, moral facts describe how we should act in certain situations. In other words, moral facts refer to an objective truth about the moral status of an action, principle, or state of affairs. For example, if I were to say, "It is morally wrong to kill others for pleasure," I would be describing a moral fact. To prove a moral fact, I need to provide a factual sentence that exists independently of any beliefs, biases, and opinions. The way I establish that this fact is indeed a moral fact would have to also be independent of any beliefs, biases, and opinions.

First, I argue that moral facts are rooted in objectivity since, as humans, we are volitional beings who use our volition to preserve and continue our existence. In doing so, we demonstrate that our existence is valuable (so every time we choose to eat, we prove that we value our life - and your life is the only thing that enables you to hold values). Thus, if values exist at all, then the volitional consciousness that creates the values is itself a value; this value is a priori, thus unprejudiced and unbiased. In summary, so long as we live, we hold ourselves as at least one of our values; thus, a universal value exists before all other values and experiences.

Additionally, we live in a cause-and-effect world. Given some set of values, the effects of a decision will be either in line with said values or not. Some actions will necessarily not achieve value, regardless of the actor's opinion, and others will do so, again, regardless of the actor's opinion. This overarching value, I argue, is morality, as this inherent value, which we hold for our existence, extends beyond mere self-preservation. Morality has historically transcended individual preferences and encompasses a broader understanding of what is valuable in our lives. As such, it is the only value that could fit this criterion as applying this universal value to our actions and decisions in the world. Conversely, actions that deviate from or contradict this fundamental value can be deemed morally wrong. In a predictable cause-and-effect world, the consequences of our

actions can be evaluated in relation to this underlying value. Some actions inherently contribute to realizing value, while others inevitably undermine it. This objective evaluation of actions based on their alignment with the fundamental value of existence forms the basis of moral judgment.

Of course, morality does not provide empirical facts like scientific research does, so it is hard to find objectivity in complex moral actions or dilemmas. Nevertheless, this does not mean there is no objectivity; rather, when it comes to morality, it requires a different method of understanding the facts. Disagreeing about morality does not necessitate that there is no objectivity; there are disagreements in every discipline. For example, people disagree about physics, but nobody considers it a reason to doubt it. Most people—or everyone—could be wrong, and the same applies to morality. Nevertheless, we agree that there is objectivity in those disciplines, which I argue we should be attributing to morality, too. Consider mathematics: it utilizes logical reasoning to prove mathematical theorems. This logical argument cannot be quantified empirically, as mathematical elements like numbers are not physically observable in the natural world. Nevertheless, they are essential for the advancement of science and are objective in that there are mathematical truths that remain true regardless of our opinions on them. Therefore, the assertion that objective facts must be scientific or derived from God is a metaphysical blunder.

But how can we determine its nature if morality should not be validated via scientific inquiry? What tool is used to locate an OMR? That tool is us; we are the instruments that establish it. For instance, imagine a situation where someone close to you refers to you as repugnant and crude. If you highly regard their viewpoint, you would almost certainly feel upset by being branded as such. This emotion provides proof of a different kind; not measured with a ruler or sample, but still valid. Just as I cannot declare a vaccine safe using a ruler, using empirical research to measure morality would incorrectly measure the facts. We are the measures of morality, and we can establish what is moral or not through shared feelings and experiences, such as pain. When deciding what to believe, we should only start with the premises we are most confident in, and I would argue that we are all confident in experiencing pain.

By applying Swinburne's logic and stacking method to ascertain the nature of moral facts and non-facts, I will establish simpler claims and establish them as factual, rather than focusing on complex, moral dilemmas. If we have confidence in at least one moral proposition, then this proves the existence of moral facts. Take the claim that "it is objectively wrong to torture your infant son to death for fun." This claim seems as secure as I can see with my own eyes. It seems more so: if

I somehow became convinced that either I was hallucinating or torturing my infant son to death for fun was right, I would immediately assume I was hallucinating. This claim certainly seems more secure than "if people disagree about morality, there might be no right answer." I have got no reason to abandon a premise I am very secure in for a premise that just seems plausible. Therefore, I would argue that this is an example of one moral claim independent of my beliefs, biases, and opinions, especially since it is one everyone believes.

Like all moral facts, that moral fact seems to surround one feeling: pain. Pain (both in the physical and emotional sense) is a universal experience that we are all familiar with and have experienced and could serve as a basis for establishing moral facts. For example, it is a universal truth that we all experience pain when insulted by someone whose opinions we value. Thus, we should refrain from insulting others. Of course, there are certain scenarios where this is a slippery slope. However, areas of apparent moral disagreement, such as the arguments over gay marriage, often rest on a disagreement about non-moral matters: for instance, whether same-sex parenting causes children psychological distress. By establishing that, at bedrock, it is generally good to avoid insulting others, we strive to avoid experiencing negative emotions like aggression, so we behave in a manner that minimizes the occurrence of these emotions and ensures others do not experience them. Although the occurrences of these items may be ambiguous in certain instances, it does not imply that there is no solution. As a society, we will gradually determine the most advantageous scenario for us, as we have with other issues, and universally agree on fundamental principles. Just as how slavery was formerly considered acceptable despite the presence of religion, over time, it was gradually eliminated, leading to the institution of slavery being outlawed and no longer widely practiced as it was in the past. We have found the objectivity: enslaving people is morally reprehensible, and this will soon apply to other moral maxims under debate.

SECTION IV: THE C.A. COUNTERARGUMENTS

The C.A. may have some counterarguments to my claims, which I will consolidate into two main points:

1. The absence of God leads to *moral relativism*, wherein moral judgments are contingent upon individual or group standards, lacking an objective arbiter; for instance, observing diverse cultural norms worldwide underscores this variance in moral perceptions.

2. Concerning the Euthyphro dilemma and the D.C.T., God, in creating us, designed us to thrive when we act in accordance with his inherent nature, goodness. For example, actions like lying or harming others go against our nature and do not lead to fulfillment. So, God's commands are not arbitrary because goodness is an inherent aspect of God's character, and all orders he issues align with this inherent goodness.

To address the first claim, it is wrong for the C.A. to claim that there cannot be an OMR because moral codes are tailored to or by a particular culture. Regardless of the subjective cultural variations, the identical objective moral principle is applicable. Most more complex moral judgments are derived from a few basic ones, with components that vary with the material conditions of different societies. However, the fact that different societies make different moral judgments does not prove relativism. To prove their position, relativists must dig down to the fundamental moral judgments in every society and then show that societies do not share these judgments, which has not been done. It cannot be denied that all societies share certain values necessary for any society to function (for example, no lying, promise-keeping, nurturing children), and as such, each culture is objective, even if there are differences at face in every culture.

For example, in one culture, it is more appropriate to meet someone with a bow; in another society, a handshake is the customary greeting. Although there are subjective cultural variations in greeting customs, the fundamental moral value of demonstrating respect and recognition toward people stays constant in both cultures. Moral virtues like respect, compassion, and love are generally acknowledged yet can be interpreted subjectively. Although we understand the importance of respecting and being nice to one another, we often achieve these goals through various methods that ultimately yield similar outcomes.

Now, in rebuttal, one might list the wickedness of our world and the evil that has been created. Many people, like Stalin and Hitler, have sought to construct their morality and impose it on others simply based on their subjective desires, and as such, they have created a new society. With that, it could be argued that we can clearly create a morality that is not independent of biases but benefits a certain group of people. I agree with that; however, we can now see that what they developed and did was morally wrong, demonstrating our objectivity, how we serve as tools for discovering objective moral truths, and why an OMR exists. The communities and moral laws they all attempted to construct resulted in inhumanity, which is why they collapsed. It is important to

clarify that while foundational moral facts are objective, they are fitted to, or shaped by, the culture and period in which they are incarnated. However, it has never been true in any society at any time that you may go out and kill someone without provocation or consequence. The only way to avoid punishment, no matter how severe, is to shun the court system (no matter what it says). Therefore, the same objective moral principle applies regardless of subjective cultural variations.

Additionally, the disagreement of moral differences in every society is often exaggerated; there is a core set of moral norms that almost all humans accept (or rather, should accept). For humans to live together in peace, we need to establish moral facts, such as not killing, stealing, inflicting pain gratuitously, telling the truth, keeping our promises, reciprocating acts of kindness, etc. The number of core norms is small, but they govern most of our transactions with other humans, and why they have remained consistent for so long. This is why we see these norms in all functioning human societies, past and present. Any community lacking these norms could not survive for long, which asserts my claim that there must be certain moral facts for us to even function. This shared core of moral norms represents the common heritage of civilized human society.

To address the second claim, the C.A. must clarify the factors that shape God's inherent character and provide reasoning for the assertion that there is intrinsic worth in human people and their nature, which affects God's decisions. Essentially, they need to determine if God selected his own character qualities. The next logical step in this dialogue would be to inquire about the foundation for his belief in his intrinsic goodness. Was there a specific reason for this? If a reason exists, the C.A. must posit that reason decides the ultimate moral norm rather than God himself. This implies that morality is outside God's influence, contradicting our established concept of God. If the C.A. asserts there was no reason, morality is once more considered arbitrary. If the C.A. asserts that God did not select his own character attributes, it implies that immutable rules bind God and lack autonomy. Finally, if God lacks control over his intrinsic character, it contradicts our concept of God and implies that he is just following some sort of predetermined programming, like a robot, rendering morality arbitrary.

Conclusion

I hope to have demonstrated that the existence of objective morality is possible without the presence of a deity; it only requires a non-empirical method of measurement. Of course, defining

what is moral and immoral is much more difficult and complex than determining an empirical or mathematical fact. Nevertheless, there must be a reason why there is still a debate about morality; why this issue has been debated for so long, and why this paper has been written. There are no serious philosophical concerns or debates regarding whether apples are better than bananas, as everyone understands the answer to be purely subjective. Therefore, morality's controversial nature is a reason to think that objective truths are at stake. Discovering God does not enhance morality, and rejecting God does not diminish morality. Regardless of religious beliefs, it is universally understood that harming an innocent child for fun, for example, is immoral, while showing respect to others is moral.

BIBLIOGRAPHY

Swinburne, Richard. (1979) *The Existence of God* (New York: Oxford University Press).

CONTRIBUTORS

AbdulRahman Alzarouni, B.A. Philosophy, The Ohio State University. Epistemology, Metaphysics, Philosophy of Language, Philosophy of Mind. AbdulRahman intends to apply to PhD programs in philosophy.

Christopher Caldwell, B.A. Kinesiology, Occidental College. Philosophical and Mathematical Logic, Philosophy of Language, and Philosophy of Science. Chris intends to apply to PhD programs in philosophy.

Albert Cardenas, B.A. Philosophy, California State University Los Angeles, Critical Theory, Existentialism. Albert intends to apply to M.A. programs in philosophy.

Arwa Hammad, Pursuing B.A. Philosophy and B.S. Political Science, California State University Los Angeles. Philosophy of Language, Ethics, Jurisprudence, Political Philosophy. Arwa will apply to joint-MA and J.D. programs in philosophy after completion of her B.A. and her B.S.

Alexis Valenzuela Haro, B.A. Business Economics, University of California Merced. Metaphysics, Phenomenology, German Idealism. Alexis intends to continue pursuing his philosophy education at the PhD level.

Colleen Hieber, B.A. Religious Studies, California State University Chico.

Ziang Li, B.A. Philosophy, University of Chinese Academy of Social Sciences. Philosophy of Perception, Causation, Epistemic Luck, Phenomenology. Ziang intends to apply to PhD programs in philosophy.

Edie Markovich, B.A. Philosophy (minor in Classics), California State University Los Angeles. Ancient Philosophy, Metaphysics. Edie intends to nourish her love of wisdom forevermore.

Cyprus Marques, B.A. Philosophy and Psychology, University of Miami. Feminist Philosophy, Social and Political Philosophy, Philosophy of Gender. Cyprus intends to apply to PhD programs in philosophy.

Theo Morgan-Arnold, B.A. Sociology, University of California Los Angeles. 19th Century Aesthetics, Political Philosophy. Theo intends to apply to PhD programs in philosophy.

Nate Miller, B.A. Communication, Bowling Green State University. Applied Ethics. Nate intends to apply to PhD programs in philosophy.

Eric Phipps, B.A. Philosophy, California State University Los Angeles. Individuation, Ethics, Technology, and Social Political Philosophy. Eric intends to apply to PhD programs in philosophy within the next couple of years.

Austin Straley, B.A. Philosophy, California State University Long Beach. Metaphysics, Philosophy of Language, Social Philosophy. Austin intends to apply to PhD programs in Philosophy.

Thomas Takasugi, B.A. Philosophy, Economics, University of California Davis. Philosophy of Language, Metaphysics, Philosophical Logic. Thomas intends to apply to PhD programs in philosophy.

Anthony Ulibarri, B.A. Philosophy, Arizona State University. Metaphysics, Philosophy of Religion, Philosophy of Mind. Anthony intends to apply to PhD programs in philosophy upon completion of his M.A. degree.

Joe Vargas, A.A. Architecture, Pasadena City College. Philosophy of Mind, Aesthetics, Phenomenology. Joe writes software and intends to apply to M.A. 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