

# **Moral Agents in a Moral World: A New Account of Moral Realism and Moral Perception**

Submitted for the Degree of Doctor of Philosophy  
Department of Philosophy, School of Humanities

Lanell Maria Mason  
May 2021

Declaration:

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Lanell Maria Mason

*For my father,  
David James Andona (1960-2017)*

# Contents

<i>Acknowledgements</i>	v
<i>Abstract</i>	vi
<i>i. Introduction</i>	vii
<i>i.i Mapping the Thesis—Part I</i>	viii
<i>i.ii Mapping the Thesis—Part II</i>	x
<i>i.iii The Two “Oughts”</i>	xi
<b>PART I: THE BASIC METAPHYSICAL FRAMEWORK</b>	<b>1</b>
<b>1. A Case for Substance Causation</b>	<b>2</b>
<i>1.1 Introduction</i>	<i>2</i>
<i>1.2 Humean Causation</i>	<i>4</i>
<i>1.2.1 Contemporary Regularity Theory</i>	<i>7</i>
<i>1.2.2 Counterfactual Causation</i>	<i>9</i>
<i>1.2.3 Probabilistic Causation</i>	<i>13</i>
<i>1.3 Substance Causation</i>	<i>16</i>
<i>1.3.1 Lowe’s View</i>	<i>17</i>
<i>1.3.2 Problems with Non-Causal Power</i>	<i>23</i>
<i>1.4 Making Sense of an Ontology of Powers</i>	<i>31</i>
<i>1.4.1 Power and Qualitative Properties</i>	<i>32</i>
<i>1.4.2 The Manifestation of Powers</i>	<i>34</i>
<i>1.4.3 Arriving at Causal Understanding</i>	<i>37</i>
<b>2. The Metaphysics of Human Agency</b>	<b>41</b>
<i>2.1 Introduction</i>	<i>41</i>
<i>2.2 Reasons Internalism</i>	<i>42</i>
<i>2.2.1 O’Connor’s Reasons Internalism</i>	<i>46</i>
<i>2.2.2 Problems with O’Connor’s View and Internalism in General</i>	<i>51</i>
<i>2.3 Reasons Externalism</i>	<i>55</i>
<i>2.4 Motivating States of Affairs</i>	<i>62</i>
<i>2.5 The Current Picture</i>	<i>68</i>
<b>3. The Reasonableness of Qualia</b>	<b>70</b>
<i>3.1 Introduction</i>	<i>70</i>
<i>3.2 Defining Qualia</i>	<i>71</i>
<i>3.3 The Ineffectiveness of the Qualia Debate</i>	<i>72</i>
<i>3.3.1 Tactics for Dismissing Introspective Evidence</i>	<i>75</i>
<i>3.3.1.1 Collapsing First-Person to the Third Person</i>	<i>75</i>
<i>3.3.1.2 Negating the First-Person</i>	<i>80</i>
<i>3.3.1.3 Redefining the Content of First-Personal Experience</i>	<i>84</i>
<i>3.3.2 Assessment</i>	<i>92</i>
<i>3.4 Positive Argument for the Existence of Qualia</i>	<i>94</i>
<b>4. Mind and Body Interaction</b>	<b>100</b>
<i>4.1 Introduction</i>	<i>100</i>
<i>4.2 Non-Reductive Physicalism</i>	<i>101</i>

4.3 <i>Panpsychism</i>	105
4.4 <i>Problems with Panpsychism</i>	107
4.5 <i>Idealism</i>	110
4.6 <i>Objections to Idealism</i>	114
4.7 <i>Substance Dualism Revisited</i>	122
4.8 <i>Conclusion</i>	127
<b>PART II: MORAL METAPHYSICS, KNOWLEDGE, AND PERCEPTION</b>	<b>129</b>
5. The Question of Moral Perception	130
5.1 <i>Introduction</i>	130
5.2 <i>Harman's Cat</i>	131
5.3 <i>Robert Audi's Theory of Moral Perception</i>	133
5.4 <i>A New Theory of Moral Perception</i>	138
6. The Metaphysics of Relations and Moral Relations	148
6.1 <i>Introduction</i>	148
6.2 <i>Internal Versus External Relations</i>	149
6.3 <i>The Phenomenal Experience of Relations</i>	157
6.4 <i>The Metaphysics of Moral Relations</i>	161
6.4.1 <i>Final Causality</i>	165
6.4.2 <i>Moral Causation</i>	168
7. Non-Conceptual Content and Perceptual Knowledge	175
7.1 <i>Introduction</i>	175
7.2 <i>Conceptual and Non-Conceptual Content</i>	178
7.3 <i>Justification</i>	188
7.3.1 <i>Recognitional Skill Theory</i>	188
7.3.2 <i>Justification Through Perceptual Content</i>	194
7.4 <i>Non-Conceptual Content and Justified Moral Perceptual Beliefs</i>	202
<i>Conclusion</i>	206
<i>Bibliography</i>	208

# Acknowledgements

I would like to thank my primary supervisor, David S. Oderberg, for his keen criticisms and invaluable feedback. His dedication at challenging my ideas has been vital to the development of what is to be found in these pages. I also would like to thank my secondary supervisor, Luke Elson, whose advice has helped me grow as an academic, a writer, and a presenter. Also, a thank you to my fellow postgraduate researchers at University of Reading's philosophy department for their consistent eagerness to interact with my work.

My thanks also go to the faculty of Talbot School of Theology's MA Philosophy program—namely Greg Ganssle, Doug Geivett, David Horner, and J.P. Moreland—in La Mirada, California. It was through these faithful stewards of knowledge that the philosophical foundations for this thesis were laid.

I would also like to extend thanks to my friend and fellow academic journeyer Emily Stevens, who is a constant listening ear. I also thank my husband, Phil Mason, who never fails to have faith in my philosophical ability. Without his love and support, this thesis would not be possible. I also give a little thanks to my daughter, Evelyn, who keeps me grounded by reminding me of what is important. And I'd like to give an extra special thanks to my parents, David and Teresa, who have, ever since I was a small child, never doubted my ability to achieve academically and have never let me doubt it either.

# Abstract

The purpose of this thesis is to provide a metaphysic for moral realism and moral perception. This thesis is in two parts. The first is concerned with basic ontology. I begin in chapter 1 with an analysis of causation, demonstrating that substance theory is superior to Humeanism at accounting for our observations; thus I defend a substance ontology. In chapter 2, I address human agency, demonstrating that reasons internalism does not allow for incompatibilist freedom; hence, I affirm reasons are states of affairs. I move on to qualia in chapter 3, demonstrating that arguments both for and against qualia realism involve question begging. In response, I argue that qualia realism is just as reasonable as anti-realism, and affirm realism. Given this, I address the mind-body problem in chapter 4, demonstrating the failure of physicalists to account for mental causation. I conclude that panpsychism, idealism, and substance dualism are equal reasonable alternatives: each allows for intrinsic properties, which are integral to substance ontology.

The second part builds a moral ontology, along with an account of moral perceptual knowledge, compatible with the above basic ontology. In chapter 5, I reject moral properties as monadic since this would bar mapping moral concepts onto mind-independent reality. Instead, I argue that moral properties are *relations*. More specifically, in chapter 6, I make the case that they are *features* of relations, relations themselves being the total complex states of affairs between one or more substances. This allows moral properties to have their own non-reducible phenomenal quality. To account for moral normativity, I argue that moral features are actually a class of causal features, and briefly argue that final causes determine the normative causal trajectory of all substances. In chapter 7, I conclude this thesis with addressing possible objections to my theory of perception and perceptual knowledge.

## i. Introduction

The purpose of this thesis is to reconnect ethics and metaphysics in a serious way—a way that has been largely lost in post-enlightenment ethical and metaethical discussion—and to provide a careful and comprehensive ontological account of morality. The theory defended here holds that moral properties are a certain class of ordinary<sup>1</sup> properties. Moreover, it will be argued that moral properties possess their own phenomenology so as to be perceivable to human subjects via ordinary sense perception.

It may be helpful to start by situating this thesis within the broader contemporary discussion of moral metaphysics. We can begin by asking whether my theory is cognitivist or non-cognitivist. Because my theory fixes moral properties externally, it is cognitivist: moral judgments are either true or false and what makes them so is situated in the external world. What is being offered here therefore is a new variant of moral realism. However, the theory developed does not fit cleanly on one side or the other of the naturalism versus non-naturalism distinction between moral realist theories. This is because, although I hold moral properties to be a certain class of ordinary properties, these properties are not the objects of scientific inquiry, at least in any recognizable sense in which science is contemporarily practiced. Yet, neither are moral properties *sui generis*.<sup>2</sup> They are ordinary features of the mind-independent world. This thesis therefore entails a rejection of both emotivism and constructivism as well as realist theories that ground moral truths via an internal appeal to rationality or practical reason.

---

<sup>1</sup> By ordinary properties, I mean concrete, natural, non-*sui generis* properties, such as color and shape properties.

<sup>2</sup> See Russ Shafer-Landau, *Moral Realism: A Defense* (Oxford: Oxford University Press, 2003), 59, 66, for a fuller description of the definition of ethical naturalism and non-naturalism being referred to here.

Moreover, accounting for the moral imperative in this thesis is not as clear-cut a matter as contemporary normative reasons internalism versus externalism debate seems to suggest. Instead, the theory defended here holds that moral obligation and the imperative comes from a mixture of external and internal considerations. The former concerns the nature of external ontology, while the latter concerns the internal ontology of the agent herself. Hence, this thesis will not only address moral properties, but also the ontology of agency, mind, and perception. It is through these that the “ought” of “something ought to be done” is externally grounded, and that the “ought” of “the agent ought to do something about it” is internally grounded. Keeping these two moral grounds distinct helps clarify what otherwise might seem like a winding discussion of otherwise disconnected metaphysical and epistemological topics.

### **i.i Mapping the Thesis—Part I: The Basic Metaphysical Framework**

Part I of this thesis builds the basic metaphysical framework onto which a moral ontology may be attached, and I start constructing this framework from a place of primary importance to any ontology: causation. In chapter 1, I defend a neo-Aristotelian view of causation—namely, E. J. Lowe’s substance causation—against Humean regularity theory and its contemporary iterations. Starting with a survey of Humean accounts of causation, I argue that Lowe’s substance causation is far superior to any Humean theory at accounting for a single instance of causal activity. However, Lowe’s theory is not adopted wholesale. Specifically, his notion of a noncausal power is rejected. Instead, an alternative analysis of the role of active and passive powers will be given and, moreover, an ontological analysis of power properties themselves will be provided and the notion of final causes will be introduced. The importance of this chapter is in its demonstration of the *existence of substances*, along with the kind of causal activity that an ontology of substances entails.

In chapter 2, I continue the causal analysis and address the issue of human beings as causal agents. The purpose of this chapter is to preserve *incompatibilist freedom* (or what some philosophers call “libertarian free will”) in an account of human agency. This is done by way of an analysis of motivating reasons, with the focus being on the causal role of desires. I argue that any theory that takes motivating reasons to be desires—or any other internal state—is deterministic and inconsistent with incompatibilist human freedom. Instead, I borrow from Jonathan Dancy in holding that reasons for actions are states of affairs. This allows the human agent to perform rational action and yet not have the reason for that action to be causally deterministic.

I then move onto the question of the existence of qualia in chapter 3. In this chapter, I contend that arguments both for and against qualia realism involve question begging. This is because philosophers must decide what to do with introspective evidence—whether to grant it epistemic authority—before they begin argumentation. However, deciding what to do with introspective evidence requires evaluating it, and evaluating it prior to argumentation begs the question in one direction or another depending on one’s evaluation. In response, I argue, through an analysis of eliminativist theories, that qualia anti-realism is no more reasonable than qualia realism and therefore there is nothing barring us from adopting realism. I also provide a positive argument for the existence of qualia (although this argument assumes the value of introspective evidence). The function that this chapter performs in the thesis is twofold. First, it demonstrates the existence of qualia, which chapter 4 will later presume. Second, it provides the metaphysical framework in which *phenomenal properties* exist, which will be important to part II’s discussion of moral phenomenology.

Chapter 4 deals with the mind-body problem. Given the assumption of the existence of qualia along with an assumption of a physical external world, I survey possible solutions to the mind-body problem. In the end, however, I commit to no one theory of mind, since

each have their own grave difficulties. I begin by dismissing substance dualism on the basis that it violates causal closure. Non-reductive physicalism is then rejected on the basis that it bars real mental causation. I then give a fair hearing to panpsychism and idealism. Both theories, however, have their own seemingly insurmountable problems. Lastly, I reconsider substance dualism and various dualist attempts to solve the interaction problem. I conclude that any theory which allows for the existence of intrinsic (to the substance) properties can be reasonably adopted. These would be panpsychism, idealism, or substance dualism. The reason why intrinsic properties are important is because their existence allows not only for *real powers* but also *mental causation*, thereby upholding both the substance ontology and incompatibilist human freedom demonstrated in chapters 1 and 2.

### **i.ii Mapping the Thesis—Part II: Moral Metaphysics, Knowledge, and Perception**

Using the framework constructed in part I, part II develops a theory of moral perceptual knowledge and a corresponding moral ontology. Moral truthmakers are situated in the mind-independent world, and an account of how moral agents can come to know moral truths is provided. The theory developed here does not deny that practical reasoning can discover moral truth. What is denied however is that practical reasoning is the sole avenue to moral truth and/or the ground of moral truth.

I introduce part II in chapter 5 with the question of whether moral concepts map onto mind-independent reality. By way of analysis of Robert Audi's theory of moral perception, I reject moral properties as monadic and supervenient since taking them to be so bars us from escaping a subject's internal state and attaching moral concepts onto external entities. Instead, I argue we should take moral properties to be *relational* properties between substances. Moreover, I account for the failure of some subjects to recognize moral properties, not in terms of their failure to qualitatively experience moral properties, but because of their failure to possess the relevant moral concepts.

Chapter 6 deals specifically with developing a moral metaphysics that allows for moral perception. I begin by demonstrating that relations—moral or otherwise—have their own phenomenology above and beyond that of the monadic properties involved in relations. Given this, I develop a novel theory of relations which holds that relational propositions do not actually pick out relations, but only *features* of relations. Relations themselves are the total complex states of affairs between one or more substances that provides the truthmaker for any number of relational propositions concerning that given state of affairs. Given this, all relations are non-reducible to monadic properties, thereby allowing them to have their own *non-reducible* (to monadic properties) *phenomenology*. I then move to argue that certain kinds of causal features of relations are what make them moral. In defense of this, I propose that final causes determine the normative causal trajectory of all substances, and that a moral agent acting to ensure that sentient substances causally trend towards their final causes is the definition of right moral action.

Chapter 7 deals with epistemological problems that arise in my account of perception developed in chapters 5 and 6. That is to say, in order to account for the fact that some subjects make correct moral judgments based upon moral perception while others fail to, I assume both the existence of non-conceptual content and *non-conceptual knowledge*. The existence of both non-conceptual content and non-conceptual knowledge is highly controversial in epistemology, however, so chapter 7 consists in a defense of these, in particular the latter.

### **i.iii The Two “Oughts”**

Recall my claim that the moral imperative is grounded both externally and internally, through the “ought” of “something ought to be done” and the “ought” of “the agent ought to do something about it.” Now we can discover how each of the concepts demonstrated in this thesis—the existence of substances, incompatibilist freedom, phenomenal properties, real

intrinsic powers, mental causation, moral (features of) relations, moral phenomenology, and non-conceptual knowledge—contributes to this grounding.

In terms of “something ought to be done,” substances and their final causes determine the quality of any moral relation into which said substances enter. If a substance finds itself in a situation where another substance’s causal activity, or even its own causal activity, interrupts its causal trajectory towards its final cause, then *something ought to be done*. We see this *something ought to be done* as consisting in causal relations, specifically causal relations that interrupt one or more substances movement towards its final cause.

Of course, this doesn’t just pertain to any and all kinds of substances. There’s nothing immoral about oxygen causing steel to rust *per se*. What is necessary to make a situation moral is the presence of a moral agent—the moral agent being the human substance. Moreover, what also seems necessary is the presence of at least one organic and significantly sentient substance whose teleological trending is at risk of being interrupted.<sup>3</sup> The presence of the moral agent himself fulfills this requirement, and thus, at the very least, a moral agent has an obligation to himself.

But what makes it the case that “the agent ought to do something about it”? The fact that *something ought to be done* doesn’t necessarily ground the notion that the agent specifically has the obligation. What does ground the agent’s obligation, however, is that the agent has both the causal ability to *do something about it* and the capacity to know that *something ought to be done*.

---

<sup>3</sup> It may be perhaps that the moral agent has a moral obligation to all organic substances regardless of sentience or non-sentience, or it may be perhaps that the moral agent’s moral obligation comes in degrees in accordance with how much sentience an organic substance possesses. Working out a specific theory in this area is beyond the scope of this thesis. What is important for present purposes is the notion that moral agents have moral obligations towards other substances, but not *all* other substances. Clearly moral agents have obligations towards other moral agents, and clearly moral agents do not have obligations towards inorganic substances. This doesn’t mean that things like destroying statues or defacing natural inorganic beauty is always morally neutral. But these things are only wrong insofar as they affect the teleological movement of the agent himself, other moral agents, or lower sentient substances.

The causal ability comes from the agent being a substance and thereby possessing intrinsic causal powers. However, it is not enough just to possess causal power unless one is able to direct it towards certain goals. Therefore, the fact that the human agent has incompatibilist freedom—meaning that he possesses the ability to internally originate external physical causal chains and rationally direct them—means that the human agent can be a morally responsive agent.

However, this is not all that is necessary. What also is needed is the capacity to know that *something ought to be done*. Included within this idea is the capacity for the agent to know that he is able to make a causal difference. This knowledge can come through practical reasoning, but it also can come immediately through perceptual knowledge. The fact that this is a capacity is important: all normally functioning agents possess the ability to see moral phenomena, and moral formation involves mapping concepts onto these phenomena.

Thus, that *something ought to be done* comes by way of the *existence of substances* with *real causal powers*, including final causes. That *the agent ought to do something about it* comes by way of *incompatibilist freedom, real mental causation, and perceptual moral knowledge* gained conceptually and *non-conceptually* via the experience of *moral phenomenal properties*. This is the theory of moral realism that will be developed over the course of this thesis.

PART I:

THE BASIC METAPHYSICAL FRAMEWORK

# 1. A Case for Substance Causation

## 1.1 Introduction

Debates carried out by those investigating the fundamental ontology of causation have historically revolved around the question of what properties, entities, or relations are primary in the occurrence of causal change.<sup>1</sup> Those theorists who have held to *event* views of causation maintain that primacy belongs to the causal event itself, this event being comprised of a relation between what would be considered cause and effect events. Critics of event views, however, argue that the notion of the primitiveness of events has left too much unexplained in terms of *how* causes generate the effects that they do, that is, why particular events are followed by the events that they are followed by, and how we can explanatorily establish a cause and effect relationship between events to begin with. In response, some event causalists have proposed various means by which we can give an account of an effect as rightly being a product of an event playing a causal role. Such attempts have included David Lewis' notion of events having counterfactual dependence on other events, or Timothy O'Connor's notion of effects coming about due to a cause's probabilistic dispositionality. Counter to the event views, others have rejected causal primacy holding extrinsically between causes and effects—such as through correlative regularity—and have argued that true causal agency lies intrinsically within the objects that are involved in causal relationships themselves. Such are what I call the *powers* views of causation,<sup>2</sup> with various versions

---

<sup>1</sup> See also E. J. Lowe, "Personal Agency," *Royal Institute of Philosophy Supplement* 53 (2003): 211, where Lowe puts this in terms of what "ontological category or categories the items related by causal relations should be conceived to belong."

<sup>2</sup> To clarify: I am utilizing the terms *powers* and *dispositions* interchangeably, as it fits with the ontology that I develop in this chapter, and for simplicity of discussion I will be conflating *powers* and *dispositional* views of causation in order to distinguish them from *event* views of causation. However, I recognize that not all philosophers do use these terms interchangeably but rather make a distinction between *powers* and *dispositions* in their ontologies. One such philosopher is Alexander Bird. He holds *disposition* to be a term used in macro level objectual dispositional predication—a term which carries no real ontological commitments. On the other hand, *powers* for Bird are properties which are essentially dispositional, and when used in causal predication, are indicative of necessary connections between objects. For Bird, it is questionable

holding that the powers belong simply to the objects that possess them or fundamentally to the properties that said objects bear. Thus, causal theorists have historically fallen into one of two wide yet internally diverse camps: either they hold to a form of *event causation* that places causal primacy on the event level of causation, or they hold to a view of *powers* that places causal primacy at the object level.

In this chapter, I will be defending a powers view of causation. More specifically, I will be proposing a *neo-Aristotelian* view of powers, as opposed to any type of Humean view of causal regularity. In order to motivate a neo-Aristotelian view specifically, I will begin this chapter by analyzing Hume's event regularity theory, as well as providing a brief analysis of a variety of recent *event* causation theories that attempt to mitigate the failures of Hume's theory to explain why certain events tend to be followed by the events that tend to follow them. I will reject Humean-style event causation on the basis that, even though it is epistemologically useful, it does not provide a reasonable metaphysical account of causation.

Next, I will present an account of causation similar to the one that I ultimately wish to defend, that is, E. J. Lowe's substance causation theory, which adheres to a neo-Aristotelian view of powers. I will begin this section by presenting Lowe's theory as he espoused it, followed by providing argumentation for why certain aspects of Lowe's view, namely his notion of *non-causal power*, should be rejected. Last, I will analyze powers views of causation in general in order to derive a theory of powers that provides a reasonable account of a neo-Aristotelian ontology of causation.

---

that *powers* can be used in causal predication since he is skeptical that powers really have this kind of causal role. See Alexander Bird, "Overpowering: How the Powers Ontology Has Overreached Itself," *Mind* 125 (2016): 360-361 and Alexander Bird, "Fundamental Powers, Evolved Powers, and Mental Powers," *Aristotelian Society Supplementary* 92 (2018): 247-248, 253-254.

## 1.2 Humean Causation

Hume held that causation consisted in the spatial contiguity, temporal succession, and constant conjunction (that is, regular observed concurrences) of existents.<sup>3</sup> For example, we can say that, in a particular instance, water causes salt to dissolve because the water and the salt are spatially contiguous, the salt dissolves after (temporally) it comes into contact with water, and we have observed this happen to water and salt in the past. Moreover, what we perceive of as “powers” intrinsic to the objects involved in causal events or the “necessity” of causal relations is merely a product of, and internal to, the human mind.<sup>4</sup> However, as pointed out by Molnar, whether or not Hume himself held that real, but unknowable, causal connections existed mind-independently between objects is indeterminate from Hume’s own work.<sup>5</sup>

Molnar identifies Hume’s position as deriving from his splitting all objects of human inquiry into two classifying groups: “relations of ideas” and “matters of fact.” The former are known *a priori*, the latter *a posteriori*.<sup>6</sup> Since necessary causal connections cannot be naturally deduced from other known concepts (i.e., known *a priori*) or empirically observed (i.e., known *a posteriori*), then necessary causal connections as a concept cannot be applied to anything in the external world.<sup>7</sup> Thus, a Humean conception of causation is not so much concerned with the ontological as it is with the *epistemological*, that is, with providing a

---

<sup>3</sup> David Hume, *A Treatise of Human Nature*, ed. L.A. Selby-Bigge (Oxford: Clarendon Press, 1896), 75-76, 87.

<sup>4</sup> *Ibid.*, 165.

<sup>5</sup> George Molnar, *Powers: A Study in Metaphysics*, ed. Stephen Mumford (Oxford: Oxford University Press, 2003), 114-115.

<sup>6</sup> *Ibid.*, 116.

<sup>7</sup> *Ibid.*, 116-119.

*deductively* arrived at explanation for why certain events are observed to follow others (given that necessary causal connections themselves remain unobserved).

More recent Humean views of causation also generally take causal primacy, in any instance of causation, to be located at the event level as opposed to the object level of causation. That is, primary causal activity belongs to the relation between a cause and effect event, and thus is situated extrinsic to objects. This is opposed to the prime causal concern being with the objects themselves—or any intrinsic property held by the objects—involved in an instance of causation. Therefore, causation is not viewed as a concept that explains any metaphysical relations between objects, but one that defines and explains the *nomological* relations that exist between events. In other words, causal relations follow some form of principle in their instances and reoccurrences; however, there is nothing necessarily grounding these causal regularities at any level deeper than the causal occurrences at the event level itself.

Unqualified, basic Humean regularity theory is rife with problems, however. Because the prime source for causal explanation is the relation between events—and because nomological explanations require more than one instance of observation in order to provide a pattern that can be considered law-like, from which causal principles can be derived—regularity theorists have difficulty in providing adequate analyses of instances of singular causation (i.e., a single instance of a single cause and its effect). Take for instance the example of billiard ball *A* rolling into a stationary billiard ball *B*, causing *B* to be set into motion. A basic regularity theory would state that spatial and temporal proximity is indicative of one event being causally relevant to the other, and our regular observance of billiard balls smacking into each other would deem this behavior nomological, and therefore causal. However, can we supply an adequate causal explanation of this singular event without

reference to others like it? What if we lacked epistemic access to other events of this kind, would we then have to deny that this was an instance of causation?<sup>8</sup>

Closely related to this problem, and also stemming from the causal primacy of nomological relations, is the problem of determining causal relevance. Suppose at the same exact moment the billiard ball *A* hit *B*, a sound wave *C* coming from the ambient noise of the pool hall, and also from the voices of surrounding billiard players, also made contact with billiard ball *B*. Can we really say that the *A* caused *B* to move, and that *C* only had a minimal, if not null, effect? *A* and *C* had the same spatiotemporal proximity to *B*, and, considering that sound waves seemingly regularly hit billiard balls that tend to move in pool halls, we would be hard pressed to deem *A* as causally relevant while denying that *C* is. Thus, on a basic regularity theory, such contrastive explanations—that is, explanations that excludes some factors but acknowledges others—of causation are hard to come by.<sup>9</sup>

What follows is a survey of how contemporary Humean theorists have attempted to address these problems. I've identified three main approaches these Humean causal theorists take: (i) qualify regularity theory in some fashion so as to account for single instances of causation, (ii) have counterfactual causal outcomes bear on actual world causal outcomes, and (iii) utilize the probability of events occurring as causal explanations. This survey is by no means exhaustive, and the philosophical theories highlighted here are just representative samples. However, I think the weakness of their theories reflect a weakness of Humean causation in general, and thus my objections to these representative theories can be applied to Humeanism more broadly.

---

<sup>8</sup> See also Molnar, *Powers*, 118, where Molnar points out this same problem of deriving a causal explanation from a single instance of causal observation.

<sup>9</sup> See also David Lewis, "Causation," *The Journal of Philosophy* 70 (1973): 556-557, and Donald Davidson, "Causal Relations," *The Journal of Philosophy* 64 (1967): 691-703, where Lewis and Davidson also point out that unqualified Humean theory has difficulty in providing such contrastive explanations.

### 1.2.1 Contemporary Regularity Theory

Because analyzing single cases of causation does pose such problems for a regularity conception of causation, modern regularity theorists tend to make sense of causation via an analysis of causation in general, as opposed to particular instances of causation. All cases of singular causal events therefore must be analyzed in light of the *kinds* of events they are, and therefore event theorists generally focus on explicating the effects of causal *types* as opposed to token causes. An example of this approach to causation is the regularity theory of Michael Baumgartner.<sup>10</sup>

In order to resolve the problems inherent in the basic Humean view of causation, Baumgartner refuses to analyze causation in terms of a single cause for any one effect—instead, Baumgartner argues that an effect event is said to be caused by what he calls a “complex cause,” that is, the sufficient aggregation of all joint causes of an effect event under review.<sup>11</sup> Thus, with these parameters of the analysis, Baumgartner attempts to make sense of a single case of causation as follows:

*Singular Causation (SC):* An event  $a$  is a cause of an event  $b$  iff  $a$  instantiates a factor  $A$  and  $b$  instantiates a factor  $B$ , such that

- (a)  $A$  is a non-redundant part of a minimally sufficient condition  $AX$  of  $B$  and every factor contained in  $AX$  is causally relevant to  $B$  according to (V) [see below],
- (b)  $a \neq b$ , and  $a$  and  $b$  occur within the same spatiotemporal frame,
- (c) Every factor in  $X$  is instantiated coincidentally with  $a$ .<sup>12</sup>

In other words, if  $a$  is an instance of event type (or what Baumgartner calls a “factor”)<sup>13</sup>  $A$  and  $b$  is an instance of  $B$ , and if it can be shown  $AX$  (i.e., the “complex cause”)<sup>14</sup> is minimally

---

<sup>10</sup> Michael Baumgartner, “Regularity Theories Reassessed,” *Philosophia* 20 (2008): 329.

<sup>11</sup> *Ibid.*, 330.

<sup>12</sup> *Ibid.*, 347.

<sup>13</sup> *Ibid.*, 329.

<sup>14</sup> *Ibid.*, 330.

sufficient for  $B$ , then we can demonstrate that event  $a$  caused event  $b$ . Since complex causes may contain unknown and thus unaccounted for components, these are accounted for via variable  $X$ ; thus  $AX$  would denote the complete condition that is sufficient for  $B$ . If  $A$  is contained, non-redundantly, within the set of conditions  $AX$ , then  $A$  is causally relevant to  $B$ .<sup>15</sup> Further, if in the instance of the singular causal event under review, we see that  $X$  is instantiated along with  $A$  as instantiated in  $a$ , we know that  $a$  must have caused  $b$  since  $AX$  deterministically causes  $B$ . Thus, according to Baumgartner, we have demonstrated  $a$  caused  $b$  in an instance of singular causation.

Has Baumgartner assuaged the problem of singular causation? It doesn't seem so. This is because his principle of causal relevancy, which consists in a logical schema to place relational constraints on what would be relevant versus irrelevant regularities and causal conditions, leaves much to be desired. It is as follows:

(V) (a)  $A$  is *directly* causally relevant to  $B$  iff the following conditions hold:

- (1)  $A$  is a part of a minimal theory  $\Phi$  of  $B$ ,
- (2)  $A$  stays part of  $\Phi$  across all extensions of  $\Phi$ 's factor frame.

(b)  $A$  is *indirectly* causally relevant to  $B$  iff there is a sequence of factors  $Z_1, Z_2, \dots, Z_n, n \geq 3$ , such that  $A = Z_1, B = Z_n$ , and for each  $i, 1 \leq i < n$ :  $Z_i$  is directly causally relevant to  $Z_{i+1}$  in terms of (Va).<sup>16</sup>

The problem with this conception is that it relies on the notion of "factor frames." Factor frames consist in the set of known minimally sufficient conditions for any complex cause to produce any one effect (i.e., a set of the individual and/or jointly sufficient minimal conditions known to bring about the same effect). Factor frames are expanded as scientists conducting repeated observations discover previously unknown sufficient conditions for causal events.<sup>17</sup> How things determinately get included into the factor frame, however, is

---

<sup>15</sup> Ibid.

<sup>16</sup> Ibid., 342.

<sup>17</sup> Ibid., 340-341.

questionable, since it can be argued that this logical schema is supposed to be determining that for us. Instead, the contents of a factor frame must be in place before the above logical schema can even be applied, and even Baumgartner himself recognizes that this presents an epistemological problem concerning the status of minimal scientific theories prior to factor frame expansion.<sup>18</sup>

Baumgartner seems to believe that scientists determine what an event's sufficient cause is solely in terms of what has been repeatedly observed. However, this does not seem to be the case in the practice of science. Scientists determine cause and effect relationships not only based on what has been observed across similar events, but also what has been observed concerning the nature of the behavior of individual materials as they are involved in all the events they are involved in. Hence, in the case of billiard balls *A* and *B* and soundwave *C*, even though in every case of *A* and *C* coming into contact with *B*, *B* starts to roll, we would know that *C* has little to nothing to do with causing *B* to roll because we can observe *C*'s behavior in the circumstances where *A* is absent. The typical behavior of the objects involved in the event seem more salient in explaining what is causally relevant to the event than the observations of events of the same type. In any case, Baumgartner's schema fails to distinguish what is causally relevant without having to presuppose the causal relevancy of some objects beforehand, and we are left with the same problem we started out with.

### ***1.2.2 Counterfactual Causation***

A more popular event causation solution to the problem of a single instance of causation is that of David Lewis'. In his seminal 1973 article "Causation," Lewis utilizes the concept of *counterfactual dependence* in order to give an account of a singular instance of a cause and its effect. Drawing upon the considerations of possible world theory, Lewis

---

<sup>18</sup> Ibid., 341.

maintains that certain causal outcomes depend counterfactually on the causal outcomes of the same circumstance in all nearby possible worlds.

Counterfactual dependence as a concept pertains to the truth-making principles for conditional propositions under the theory of counterfactuals. Thus, the truth of the proposition  $A \Box \rightarrow C$  in the actual world depends on the truth of proposition  $C$  at all nearby worlds where  $A$  is true.<sup>19</sup> This dependence relation can be viewed as causal dependence when it comes to the analysis of singular causal events. Specifically, in causal analysis, counterfactual dependence holds between events' corresponding propositions. For example, Lewis expresses the proposition of event  $e$  occurring as  $O(e)$ , and this proposition is true at all possible worlds  $e$  has occurred. If we want to say  $c$  caused  $e$  in a singular causal instance in the actual world, the truth of this statement depends on whether  $O(c)$  and  $O(e)$  (and, by extension,  $\sim O(c)$  and  $\sim O(e)$ ) counterfactually depend on each other. If they do, then  $O(c) \Box \rightarrow O(e)$  and  $\sim O(c) \Box \rightarrow \sim O(e)$  is true, and  $e$  has causal dependence on  $c$ . This causal dependence implies  $c$  caused  $e$ .<sup>20</sup>

Lewis himself later addressed a problem inherent in a counterfactual analysis of causation, that is, the problem of preemptive causation. This problem comes to light when analyzing an event that has two would-be causes present and both are equally sufficient to bring about the event, although only one did. A theory's task is therefore to determine which of these two potential causes actually did the causing. However, as counterfactual causation stands, it cannot make a determination; if any one of the causes was absent, including the cause that actually did the causing, the effect would still be present, thereby leading to causal overdetermination.<sup>21</sup>

---

<sup>19</sup> Lewis, "Causation," 560-561.

<sup>20</sup> Ibid., 562-563.

<sup>21</sup> David Lewis, "Causation As Influence," *The Journal of Philosophy* 97 (2000): 182-183.

Lewis' proposed solution to the problem of preemptive causation is the notion of *trumping*. That is, one potential cause trumped the other cause in bringing about the effect in question.<sup>22</sup> For example, if both *X* and *Y* are present and are both sufficient causes for *Z*, but *X* is what caused *Z*, it can be said that *X* trumped *Y* in causing *Z*. However, as Lewis points out, this alone doesn't quite make sense because it implies that *X* was able to cut short *Y*'s causal chain somehow, and this cutting would be left unacceptably unexplained.<sup>23</sup> The solution he advances in response depends on the notion of modal *fragility*: slight differences between possible events make them distinct (that is, slight differences in causal events help determine *when* effect events occur and in what *manner* they occur).<sup>24</sup> Therefore, when analyzing the trumping cause, it is not that *X* cuts *Y*'s causal chain to go on and cause *Z*. What is happening is that *X* causes *Z* in a counterfactually dependent way as usual. If it was the case that *X* was not present to trump *Y*, then *Y* would have caused *Z* to occur at a different time and in a different manner.<sup>25</sup> With this solution the counterfactual dependence between *X* and *Z* is left intact without the causal cutting of *Y*.

However, this solution does not quite work with Lewis' view of counterfactual causation due to Lewis' own notion of fragility itself. As Neil McDonnell points out, it is too fragile: any change leading up to an effect would make the effect event too different, and thus arguably numerically distinct.<sup>26</sup> Therefore, effect events would have counterfactual

---

<sup>22</sup> Ibid., 183.

<sup>23</sup> Ibid., 184.

<sup>24</sup> Ibid., 185-188.

<sup>25</sup> Ibid.

<sup>26</sup> Neil McDonnell, "Events and their Counterparts," *Philosophical Studies* 173 (2016): 1303. See also Lewis, "Causation As Influence," 186-188, where Lewis specifically denies that his view leads to numerically distinct events between counterfactuals. He assumes his confinement of fragility to matters of time and manner is enough to keep events numerically the same. McDonnell, however, holds that changes in time and manner is enough to numerically change events.

dependence on things that they should not have counterfactual dependence on.<sup>27</sup> Take for example again our billiard balls and sound waves. If ball *A* caused *B* to roll in a silent room this would be a numerically distinct event than if *A* caused *B* to roll and soundwave *C* was present. However, factoring in *C* into a causal explanation if *C* is present is undesirable.

In response to this problem for counterfactual causation, McDonnell attempts to make events less modally fragile by adopting a context-sensitive theory of events combined with his own notion of event counterparts (that is, events that exhibit sameness across possible worlds).<sup>28</sup> McDonnell explains this context sensitivity as follows:

When the context triggers a counterpart relation which assigns relatively many counterparts for a given event, then that event is taken to be *robust*, and when the context triggers a counterpart relation in which the event has relatively few counterparts, then that event is taken to be fragile. So, the components of a double, or rather *fluctuating* standard are present in the [counterpart-theoretic treatment of events combined with a counterfactual theory of causation] view offered here.<sup>29</sup>

McDonnell argues that this notion of event robustness will prevent things such as soundwave *C* being considered a cause since the events where *C* occurred and where *C* did not occur will exhibit enough sameness to warrant *C* not being considered causally relevant.

There are a few problems with McDonnell's view. First, there is a huge epistemological barrier. How do we distinguish what is causally relevant? It does not seem that you can rely solely on McDonnell's view. This is because we do not have epistemic access to these counterpart events, and therefore have no idea how many there are for any given single event. Moreover, you need to presuppose what is relevant in order to provide an explanation for its relevancy utilizing McDonnell's theory. For example, in the case of Billy and Suzy being two sufficient causes involved in the event of a window breaking, McDonnell

---

<sup>27</sup> McDonnell, "Events and their Counterparts," 1303.

<sup>28</sup> *Ibid.*, 1291-1292.

<sup>29</sup> *Ibid.*, 1304.

argues that “ [i]f we take the event to be fragile enough to tell Billy and Suzy apart...but no more fragile than that requires,” then only relevant causes will be taken into consideration.<sup>30</sup> Hence, the event should be considered fragile enough to exclude Billy as a causal contributor, but robust enough to include Suzy, the relevant cause.<sup>31</sup> However, this way of determining what is causally relevant, especially given our lack of epistemic access to counterfactuals, is indicative of McDonnell *having what is required in mind even before considering what counterpart events to consider*. He’s presupposing what is causally relevant even before he gets to counterfactual considerations. Lastly, McDonnell maintains that although counterpart events exhibit some form of sameness, they still are to be maintained as numerically distinct.<sup>32</sup> If this is so, then fragility across the board still stands, and the counterfactual causalist, like the regularity theorist, cannot make sense of causal relevancy.

### ***1.2.3 Probabilistic Causation***

Probabilistic theories of causation hold that a cause increases the probability of the effect’s occurrence—thus, probabilistic causation distinguishes itself from a traditionally Humean theory in that it admits the indeterminacy of an effect event the face of what a traditional Humean would consider a sufficient cause. I do not want to spend too much time addressing probabilistic causation for a couple of reasons. First, not all probabilistic theories analyze causation via a Humean reduction to non-causal observables; some merely give an account of probability as it pertains to causation, however causation may be cashed out.<sup>33</sup> Therefore, a probabilistic theory of causation cannot always be rightly called a theory of

---

<sup>30</sup> Ibid., 1305.

<sup>31</sup> Ibid.

<sup>32</sup> Ibid., 1292-1293.

<sup>33</sup> Michael Tooley, “Probability and Causation,” in *Cause and Chance: Causation in an Indeterministic World*, ed. Phil Dowe and Paul Noordhof (New York: Routledge, 2004), 77, 82.

causation *per se* without some further qualification, nor are all probabilistic theories even Humean by design.

Second, probabilistic theory, when taken as a reductionist theory of causation, seems to falter under the same criticism as the above theories. As with other theories that place causal primacy at the event level, probabilistic causation as a causal theory has a difficult time accounting for causal relevancy, and more specifically, with providing contrastive explanations, although it does so for different reasons than regularity theory and counterfactual causation. Take for example Timothy O'Connor's argument utilizing the case of syphilitic patients contracting paresis. O'Connor points out that only syphilitic patients can contract paresis, but not all syphilitic patients do contract paresis—paresis seems to only occur in about 28% of patients with syphilis. O'Connor argues that it is the case that syphilis *causes* paresis, although it causes it *indeterministically*. He concludes that contrastive explanations are not necessary since there may not be such an explanation available to the situation; nothing is determining which patients come down with paresis.<sup>34</sup>

In response, I argue that O'Connor conflates the epistemological usefulness of these types of explanations for belief formation and justification with the ontological necessity of certain causal processes in the production of effects that these types of explanations reflect. Yes, we may not know *why* a certain patient developed paresis while another did not, and, considering the limits of modern medical science, the explanation that syphilis caused this disease is sufficient as a reasonable explanation to conclude that a patient contracted paresis because he had syphilis. But this epistemological concession does not mean the disease was really caused *indeterministically*, or at least, it is insufficient as an argument for this

---

<sup>34</sup> Timothy O'Connor, *Persons and Causes: The Metaphysics of Free Will* (Oxford: Oxford University Press, 2002), 91-93.

conclusion. It may just mean we do not know the determining factors of the disease's coming to be in the patients that it did come to be in.

Further, how do we determine causal relevancy based on this probabilistic conception? O'Connor seems to argue that it can be concluded that syphilis causes paresis because it occurs in 100% of paresis patients. But how is it determined that syphilis is causally relevant versus other considerations that also apply to 100% of the patients in question? All of them are human beings, all of them drink water on a regular basis, and all of them receive treatment for syphilis. How is it determined that these factors are not causally relevant, especially in terms of causal explanation, based on O'Connor's conception of probability theory? It seems either that we have entered some muddied causal waters, or we must claim probabilistic explanations cannot properly be considered causal explanations.

The analysis of probabilistic theory highlights a problem that has been latent throughout all of our discussion concerning causation thus far: the problem of conflating acceptable causal explanations with adequate metaphysical accounts of causation. Causal explanations are a matter of epistemology: acceptable causal explanations are belief justifying and of instrumental value. If I come into a large dark room full of people at the other end and turn on the light switch by the entrance, everyone in the room, if they had knowledge of where the light switch was located and were thinking reasonably, would attribute the light coming on to someone new coming into the room, namely, to my coming into the room. The whole event could be summarized by stating that I caused the light to come on in the room. Therefore, my coming into the room and turning on the switch is an adequate as explanation for the light coming on, and not only that, it warrants belief that I have just now come into the room.

This explanation however is not a metaphysical account of the causal processes involved in my flipping on a switch and light flooding a dark room. In fact, it makes a very

poor metaphysical explanation. What about the other objects involved in this causal process, such as the metal wiring in the walls and ceiling, and the electrical current running through them, not to mention the filaments in the light bulbs, or even the photons radiating from them? These cannot be ignored. Do we need an adequate account of these in order to accept the claim that I caused the light to come on in the room? Of course not. But we would need to factor them into the causal process if we had any desire to come away with an account of causation that had adequately reflected the full reality of the mind-independent universe.

The Humean picture, however, has ignored the question of metaphysics. Those philosophers who hold to Humean causation have found adequate epistemological explanations and repurposed them as metaphysical theories of causation. Yet, these theories leave too much of what we experience of the cause and effect relationship unexplained, with their adherents attempting to redesign them over and over again in order to fit our most common intuitions concerning the causal relevancy of objects involved in causal events. And here lies, I argue, the beginning of an adequate theory of causation: in the analysis of the objects that are involved in causal events, and in the real powers that these objects possess.

### **1.3 Substance Causation**

In order to provide an adequate metaphysical explanation, as opposed to epistemological explanation, of the causal relationships inherent in cause and effect events, I will be arguing for a neo-Aristotelian view of real causal powers inherent in the objects involved in causal relations; it is only through the acknowledgement of these powers that causation as a metaphysical reality, as opposed to an explanatory tool, makes any sense. Thus, the view I am for the most part defending is that of E. J. Lowe, which he dubbed *substance causation*. This section provides an analysis of this view. However, as will later become apparent, Lowe's theory has some weaknesses, which I will identify before developing an improved account of substance causation. That said, Lowe's view is a good

starting point for developing an accurate view of the metaphysically real causal relationships that exist between entities.

The very nature of substance causation as a view of causation, especially as Lowe presents it, requires the discussion of human persons as causal entities. However, it is not the purpose of this section to provide a thoroughgoing ontology of human persons as causal agents: this subject will only be touched upon insofar as Lowe utilizes it to frame his theory of causation. Human persons as causal entities shall be more thoroughly treated in the next chapter.

### ***1.3.1 Lowe's View***

To begin, it is important to understand what is meant by *substance* in order to discuss substance causation. According to Lowe, individual substances are concrete (as opposed to abstract) basic particulars, and includes things such as ordinary material objects (e.g., rocks, trees, billiard balls, etc.) as well as human persons.<sup>35</sup> Individual substances are *persistent* in that they are able to undergo and survive *change*.<sup>36</sup> For example, a sapling that grows into a large oak tree is still considered the same substance although it has taken on mass and changed its shape. Therefore, from what can be inferred from Lowe's definition of a substance, an object that fails to survive change can be said to no longer be the same substance, but now is a new substance or substances, the old substance having gone out of existence. A large boulder that has been broken up and split into paving stones has ceased to exist, while the several paving stones made from the constitutive material of the boulder have now come into being.

---

<sup>35</sup> E. J. Lowe, "Substance Causation, Powers, and Human Agency," in *Mental Causation and Ontology*, ed. S. C. Gibb, E. J. Lowe, and R. D. Ingthorsson (Oxford: Oxford University Press, 2013), 154.

<sup>36</sup> *Ibid.*, 154-155.

Lowe adds that individual substances have *ontological independence*, meaning they do not depend on any other entity or entities for their *identity*. This claim doesn't preclude that individual substances do depend on other entities for their *existence*, however.<sup>37</sup> Each of the brand new paving stones depend on the atoms and molecules that materially comprises it for its existence. I myself depend on the cells that make up the tissues and organs of my body for my very own existence. However, a paving stone will survive a few atoms being abraded away from its surface, and I can and do survive cellular and tissue changes. I can even undergo massive changes, such as organ and limb removal. We would not say, however, that I am a numerically different substance, that is, an entirely different person if I happened to lose one of my arms, but we would also not say that I still existed if all there is left of me is an arm. Existence and identity are distinct concepts.

Not only do individual substances have this ontological independence when it comes to identity, but they also cannot be causally inert—that is, *all* individual substances possess a power or multiple powers.<sup>38</sup> Therefore, differing *kinds* of substances are primarily differentiated by their *identity conditions* and the *power* or *powers* they possess.<sup>39</sup> In other words, how an individual substance persists through time and what changes it can be the cause of will differ from substance kind to substance kind. The conditions under which I am said to persist through time and the changes that I am able to bring about are different than that of the paving stones, and this is because the paving stones and I are different *kinds* of substances. But the paving stones themselves all share similar identity conditions and powers, because they all belong to the same substantial kind.

---

<sup>37</sup> Ibid., 154.

<sup>38</sup> Ibid.

<sup>39</sup> Ibid.

In explicating the ontology of powers themselves, Lowe makes a distinction between power *types* and power *tokens*. Token powers are those particular powers that are possessed by any particular individual substance and that substance alone.<sup>40</sup> A particular billiard ball is going to have its own particular power to roll and knock into other billiard balls; this power is, ontologically speaking, numerically different than any other power in existence. However, the token power of this particular billiard ball is of the same *type* of power shared by all other similarly constituted billiard balls.<sup>41</sup> Further, as in the case of every substances' possession of a token power, each particular billiard ball possesses its own particular power *essentially*; a particular power cannot be transferred from one billiard ball to the next, or to any other individual substance.<sup>42</sup>

All powers, regardless of their being a type or token, are differentiated by the nature of their *manifestation* (i.e., by the characteristics of the actual exercise of the power in question). There is only one manifestation type per power, that is, a power can only perform one kind of effect and none other.<sup>43</sup> However, there are *higher-order* powers that consist in the power to gain *lower-order* powers. The example Lowe uses to illustrate this point is magnetizability: this power consists in the ability to gain the power of being magnetic, itself the power to attract ferrous metal.<sup>44</sup> Moreover, because they are themselves concrete particulars (i.e., individual properties possessed by individual substances), token powers can be further differentiated by *the individual substance that possesses them* and by *the time at*

---

<sup>40</sup> Ibid., 155; E. J. Lowe, "The Will as a Rational Free Power," in *Powers and Capacities in Philosophy: The New Aristotelianism*, ed. Ruth Groff (New York: Routledge, 2013), 173.

<sup>41</sup> Lowe, "Substance Causation," 155; Lowe, "The Will," 173.

<sup>42</sup> Lowe, "Substance Causation," 155; Lowe, "The Will," 173.

<sup>43</sup> Lowe, "Substance Causation," 156.

<sup>44</sup> Ibid.

which they are possessed (i.e., the time at which they exist).<sup>45</sup> It is important to note that a token power that never manifests is still said to exist since it is nevertheless possessed by its bearer. How this can be is easily demonstrated through an example: a copper wire that fails to ever conduct electricity is still said to possess the power of electrical conductivity.<sup>46</sup> This is why Lowe argues that powers should be individuated by manifestation types as opposed to token manifestations, since it is possible that power tokens can fail to manifest during their entire existence.<sup>47</sup>

In further explicating his ontology of powers, Lowe makes categorical distinctions between different types of powers, claiming that all manifestation types are either *passive* or *active* and either *causal* or *non-causal* (see chart).<sup>48</sup> Lowe defines a *passive* power as one whose manifestation is dependent on an outside substance or substances acting upon and triggering its

	PASSIVE	ACTIVE
CAUSAL	Water's power to dissolve salt; salt's power to dissolve in water.	Gravitation
NON-CAUSAL	Ball rolling down an incline	Radium's radioactive decay; the human will

manifestation in the individual substance that possesses it.<sup>49</sup> Examples of passive powers—what Lowe also calls *liabilities*<sup>50</sup>—include salt's power to be dissolved in water and water's own power to dissolve salt. Both the substances of salt and water in this case need to be appropriately triggered by the other substance in order for the event of dissolution to occur.<sup>51</sup>

---

<sup>45</sup> Ibid., 156-157.

<sup>46</sup> Ibid., 157.

<sup>47</sup> Ibid.

<sup>48</sup> Ibid., 160. Chart adapted from Figure 6.1.

<sup>49</sup> Ibid., 153, 158-159; Lowe, "The Will," 174.

<sup>50</sup> Lowe, "Substance Causation," 159.

<sup>51</sup> Ibid., 159-160; Lowe, "The Will," 174.

An *active* power's manifestation, on the other hand, is *not caused* by any outside substance or substances acting on the possessor and trigger the manifestation of the active power. This kind of power is what Lowe calls a *spontaneous* power, due to it occurring without having to be caused. The example Lowe gives is that of a radium atom undergoing radioactive decay. The atom emits particles spontaneously, with nothing acting upon the atom and causing the particles within to be emitted.<sup>52</sup>

A *causal* power's manifestation is constituted by the possessing individual substance acting upon and effecting a change in one or more other substances.<sup>53</sup> An example causal power that is that of water's power to dissolve salt.<sup>54</sup> (This power's being causal, on top of this same power's also being passive, makes water's power to dissolve salt a *passive causal* power). On the other hand, a *non-causal* power's manifestation is *not* constituted by the possessing individual substance acting upon and bringing a change in any substance. Lowe's prime example is that of a ball rolling down an incline, since this power consists in merely translational motion.<sup>55</sup> This movement can result in a change, however: the ball could crush an object in its path on the way down the incline. The difference however is that, unlike a causal power, the non-causal power of the ball's rolling is *not constituted* by the crushing of another object.<sup>56</sup> Therefore, it seems that, according to Lowe's definitions, the main indication of whether or not a power is causal or non-causal power is whether or not its exercise is constituted by bringing about a change in another substance. In reference to the examples we have been examining, water's power to dissolve salt cannot occur without salt

---

<sup>52</sup> Lowe, "Substance Causation," 153, 159-160; Lowe, "The Will," 174.

<sup>53</sup> Lowe, "Substance Causation," 153, 157; Lowe, "The Will," 173.

<sup>54</sup> Lowe, "Substance Causation," 157; Lowe, "The Will," 173.

<sup>55</sup> Lowe, "Substance Causation," 158; Lowe, "The Will," 174.

<sup>56</sup> Lowe, "Substance Causation," 158; Lowe, "The Will," 175.

being dissolved and therefore cannot occur without another substance changing—but a ball can roll down an incline without bringing about any change in another substance.

This notion of *non-causal power*, as innocuous and mundane as a ball rolling down an incline seems, is not unimportant to Lowe's causation theory, but indeed holds a central place in it. This is because, according to Lowe, the human *will* is an *active non-causal* power: not only is the manifestation of the human will uncaused, but it also does not consist in any effects.<sup>57</sup> Moreover, human action is considered to be *substance* causation, in accordance with Lowe's theory, in that it is the human *person* who holds causal primacy concerning any effects she happens to generate through the exercise of her will.<sup>58</sup> This stands in contrast to those theorists who would claim that an agent practices agency through causing her will or intentions. For Lowe, *to will* is just to exercise agential power.<sup>59</sup>

Further, because this non-causality of the will seems to *prima facie* curtail the agent's causal ability, Lowe attempts to account for agential efficacy by appealing to the idea that one power type can result in the manifestation of another. In other words, the manifestation of a non-causal power can and does regularly result in change since it has the ability to result in another type (*active* or *passive* and *causal* or *non-causal*, whatever it may be) of power manifesting. Lowe argues that such is the example with the ball rolling down an incline having the ability to crush an object in its path, and such is the agent when she exercises the power of her will.<sup>60</sup>

Through this characterization of agential power, Lowe attempts to account for incompatibilist agential freedom, that is, the ability of the agent to originate a causal course at

---

<sup>57</sup> Lowe, "Substance Causation," 160-161; Lowe, "The Will," 174-175.

<sup>58</sup> Lowe, "Substance Causation," 162; Lowe, "The Will," 175-176.

<sup>59</sup> Lowe, "Substance Causation," 162; Lowe, "The Will," 176, 178.

<sup>60</sup> Lowe, "Substance Causation," 161; Lowe, "The Will," 175.

any given point of decision making, as opposed to agent action being causally determined by forces extrinsic to the agent's exercise of her power, including the agent's own set of beliefs and desires.<sup>61</sup> Lowe reasons that if the exercise of the will did consist in agential effects, then it could not be rightly described as free; on the other hand, if the will failed to ever result in change, then there is reason to doubt that it could not be rightly be described as the agent's power. Thus, through the notion of the will being an *active non-causal power*, Lowe wishes to retain the freedom and power of the will, and thus to account for the causal primacy and efficacy of human persons.<sup>62</sup>

### ***1.3.2 Problems with Non-Causal Power***

The notion of agent power being *active* seems intuitive enough, and acceptable enough for those philosophers willing to forego a completely deterministic view of reality. But is the notion of agential power being *non-causal* likewise acceptable? I argue that it is not, and, further, that this unacceptability stems from the internal incoherency involved in the very notion of a non-causal power.

I will begin my criticism of a non-causal power in one of the places Lowe begins an argument in favor of it, that is, with the example of the radium atom. According to Lowe, the radioactive decay of a radium atom is a good example of an *active non-causal* power in that it is clear that nothing is causing the atom to undergo this decay—to Lowe, the atom emits particles *spontaneously*—and likewise this decay does not result in any effects. Lowe seems especially pleased with this example, since, according to him, it demonstrates an occurrence of an active non-causal power in nature, and therefore its existence should be wholly acceptable even to those who are hardened naturalists.<sup>63</sup>

---

<sup>61</sup> E. J. Lowe, *Personal Agency: The Metaphysics of Mind and Action* (Oxford: Oxford University Press, 2008), 159, 175-176.

<sup>62</sup> Lowe, "Substance Causation," 163-164; Lowe, "The Will," 176-178.

<sup>63</sup> Lowe, "Substance Causation," 153.

First, to characterize this power as active in the sense that Lowe means active—that is, *spontaneous* and *undetermined*—is somewhat oversimplified. The radium atom—by the fact that it is a radium atom—possesses the *essentially* characteristic half-life of all radium atoms of the same isotope. Along with this comes a fixed and objective probability for that atom to decay within a certain timeframe.<sup>64</sup> To claim that the particle emission of decay is “uncaused” is to mischaracterize the nature of isotope decay. If it was truly as spontaneous and uncaused as Lowe wants to make it out to be, especially in light of what he wants to say about the exercise of human free will, it seems as if it would be just as likely for the atom to emit eukaryotic cells, since there would be no physical or causal constraints on what is possible;<sup>65</sup> but since we want to deny this unlikely scenario, or, at least not let it hinder our developing a theory of causation, we should not be so quick to dub this power as active in the sense that we want to say the exercise of the human will is active.<sup>66</sup> In the case of the radium, it is not so much a question of *what* will happen, but *if* and *when* it will happen—only the *if* and *when* are really undetermined, leaving particle emission unpredictable. But for a human being, we want to leave the power of *what* also open for the agent herself to determine, thus making her power truly active. In any case, the “spontaneity” of this type of micro event seems to have little bearing on the discussion of causation as witnessed in macro objects, and so should not be given such a big role in our theory of causation.<sup>67</sup>

---

<sup>64</sup> Ibid., 159.

<sup>65</sup> Ralph D. Ellis, “Three Arguments Against Causal Indeterminacy,” *Philosophia* Volume 31, Issue 3-4 (October 2004): 339-343. I am not arguing here that a human agent isn’t physically constrained in what she can do, only that she is not physically or causally constrained in what she can *determine* to do.

<sup>66</sup> Not only are eukaryotic cells possible, but, as Ellis concludes, it is just as likely for an electron to turn into a pigeon and fly away as it is for a causal event not to occur in light of causal preconditions that, barring the question of quantum indeterminacy, would be sufficient for it. See Ellis, “Three Arguments Against Causal Indeterminacy,” 343.

<sup>67</sup> See also Ellis, “Three Arguments Against Causal Indeterminacy.”

More importantly against Lowe's theory, however, is that the power of the radium atom to decay cannot rightly be called *non-causal*. Remember that being non-causal means that the manifestation of the power does not consist in bringing about change in any substance or substances. However, the emission of particles does consist in changing the nature of the atom in question in that, after an emission, the atom has in the very least less energy. In some cases, these types of emissions result in an atom changing from one element to another. Further, the particles that were emitted are no longer under the purview of the essential nature of the "mother" atom. They have become their own substances, with causal powers of their own. It seems therefore that it is wholly rational to hold that radioactive decay is not a non-causal power.

But even if my claim concerning the radium atom's causal power is met with skepticism, we can also demonstrate the unviability of non-causality using what Lowe considers to be another typical example of a non-causal power: a ball rolling down an incline. Lowe claims that the power to roll down an incline is non-causally manifested by a spherical object which possesses it. Again this would mean that this manifestation does not consist in any change occurring to any substance or substances. However, an object can be crushed as a result of the spherical object rolling over it, and this would be an instance of this power causing an effect; what's important to note for Lowe, though, is that the ball can roll *without* causing any effect, therefore making it a non-causal power.<sup>68</sup>

The problem with this account is that what we are left with is a non-causal power's operation resulting in an effect being brought about, but without an explanation of *how* an effect can indeed be caused *without causal power operating*. How exactly is it that a non-causal power's operation can be translated into causal power? In other words, what *causal*

---

<sup>68</sup> Lowe, "Substance Causation," 158.

power's manifestation resulted in the crushing of an object? Indeed, on Lowe's view, this is left a mystery. In an attempt at explanation, Lowe claims that the non-causal power is able to trigger the manifestation of another kind of power if that other power is passive and able to be triggered in the appropriate way. Thus, Lowe explains that a ball's power *to crush* an object underneath it can be manifested through the manifestation of its power *to roll* triggering its power to crush. How exactly *rolling* triggers the manifestation of *crushing*, or how any non-causal power can trigger the manifestation of another power for that matter, is left unexplained.<sup>69</sup> Nevertheless, it does seem intuitive and true to experience that a ball rolling down an incline can cause an object to be crushed under its weight, and certainly we can accept the truth of such a causal connection between rolling and crushing on theoretical grounds. But how can the truth of this be reconciled with the idea of something *non-causal* having *caused* something?

My solution involves an affirmation—yes, a spherical object can crush objects that it happens to roll over—but also a denial: what is at work in this scenario is not Lowe's notion of non-causal power, but several *causal* powers working in tandem. Remember that Lowe dubs this non-causal power the power of translational motion—the moving of the sphere from one point in space to another.<sup>70</sup> However, we can also make sense of this translational movement utilizing purely causal language without having to appeal to anything as abstruse as a non-causal power, thereby rendering any change that occurs as a result of the ball's rolling consistent with causal activity.

The spherical object, because it is a spatially extended object, is an object that possesses the property of *having mass*. All objects that have mass have causal powers entailed by mass, including the powers to be attracted and to attract other objects that also

---

<sup>69</sup> Ibid., 161.

<sup>70</sup> Ibid., 158.

possess mass. This force, which is a natural and essential feature of objects with mass, is otherwise known as *gravitational force*. The gravitational pull of the earth causally acts upon the spherical object, pulling it downwards towards the earth's center of gravity. Likewise, the sphere also pulls on the earth, but this force is negligible in comparison to that of the earth's power. Furthermore, because of it being spherical in shape, the sphere only comes in contact with the flat plane of the incline at a relatively small point in comparison to its overall surface area. As the sphere comes into contact with the incline, the surface tension of each object acting on the other causes resistance to be applied to the sphere at the relatively small contact point. Since the sphere is spherical, this easily puts the sphere into a spin if the sphere is in motion. If the earth's gravity is pulling the sphere downwards, then it is indeed in motion, and furthermore, the resistance against the earth's pull provided by the incline redirects this motion laterally. Hence, the sphere moves from one point to another as it rolls down an incline. As can be seen, it is reasonable to hold that this is not a manifestation of the sphere's non-causal power *to roll*. We can view it as several causal powers acting in tandem to move the sphere from one location to another. Characterizing this motion as the singular non-causal power *to roll* is to mistake a higher-order explanatory expression (i.e., the ball "rolling down an incline") for the fundamental first-order causal powers at work—e.g., the gravitational pull of the earth and the ball, the resistance force being applied to the ball and the incline, etc.—in the substances involved in this causal scenario.

Lowe may counter that, due to the bare fact that the ball's rolling down incline does not consist in anything more than translational motion, we can still rightly consider this rolling as a non-causal power. However, a ball's rolling down an incline *does* consist in more than just translational movement. In addition to all the effects involved in my above account, micro changes, and perhaps imperceptible macro changes, occur in the structures of both the

incline and of the ball when a ball rolls down an incline.<sup>71</sup> To deny these changes is to deny something fundamentally valuable to the substance theorist about substances in relation to their composite materials and to their interaction with other physical materials. These materials are ever changing in nature, yet the identity conditions of substances allow them to survive such changes. Even though the ball rolls down the incline seemingly with no changes, this does not mean changes have not occurred, and the substance theorist should readily admit them. Further, to perhaps characterize the difference between my account and Lowe's account as merely a difference in semantics, and thus unimportant, is to muddle the issue of developing a *metaphysical* account of causation. We need to be able to distinguish between actual first-order causal movers and the second-order expressions we utilize to describe their activity. We are attempting to develop a theory of *causation*, in the metaphysically weighty sense, not a theory of how events should be talked about.

This lack of a theoretically defensible, empirical account of non-causal power in nature puts Lowe's theory of human agency into jeopardy. Recall that Lowe believes his theory requires the exercise of the human will to be a non-causal power in order explain how the exercise of the human will is not deterministic in nature. Lowe wants the agent to be able to exercise her will without having to commit to any action that would be causally determinable, which would lead to subsequent determined effects. In essence, Lowe wants the manifestation of the will to result in only modally unnecessary actions being committed by the agent in order to preserve her incompatibilist freedom at any given point in decision making. The way he attempts to accomplish an account of incompatibilist human freedom is by postulating that the manifestation of the will does not *consist* in any action—that is, it is non-causal. However, without a plausibly held example of a non-causal power operating in

---

<sup>71</sup> I want to credit David S. Oderberg for pointing this out.

the observable, material realm, it is more difficult to accept that such a power would exist in the unobservable mental realm than Lowe would like.

A further problem for Lowe is that the agent needs to be able to affect some kind of change, or how else could an agent be considered to have agency? Lowe recognizes this need for an explanation and postulates that non-causal powers can somehow trigger other powers into manifesting. However, in light of his holding that the manifestation of the will is, indeed, a non-causal power, he fails to explain how this could possibly be. Despite what Lowe is attempting to accomplish with it, this conception of agency seems as if it could only allow the agent to produce change *per accidens*, that is, as an uncontrolled and not necessarily intended side consequence of the agent manifesting her will, and therefore not a product of any kind of agential causal determination.

Some supporters of Lowe's view may counter that this is an acceptable consequence of Lowe's theory of agency—it is better to trade an agent's ability to strictly determine what effects she causes for her ability to freely choose her actions in the incompatibilist sense. Some may even point out that our experience with exercising our powers seems to reflect this—we don't always end up with the effects that we were after when we have acted. First, in response to this objection, the unattended effects that our actions sometimes result in may not necessarily have anything to do with the ontology of our causal power and may wholly have to do with our epistemological limits: perhaps we just do not always know what actions will result in our desired effects. Since these are both plausible explanations for the same common experience, it would be unwise to accept one explanation over the other without the basis of further evidence.

Second, and more importantly, despite the *prima facie* suitability of trading agential control of agential effects for incompatibilist free will, the benefits of this tradeoff still fails to justify our accepting an ontology of non-causal powers. Indeed, if all agent effects were

fundamentally change *per accidens*, then Lowe’s entire theory of substance causation itself would be undermined, at least in the case of human agents. This is because the theory of substance causation—like all similar views that hold to the existence of *causal powers*—maintains that causal primacy belongs *intrinsically* to the entities that are involved in any given cause and effect relationship. Characteristic of the substance view, the entity doing the causing is the substance itself, as it exercises the powers that are either first-order properties or derivative from the first-order properties that the substance possesses.<sup>72</sup> The notion of non-causal powers, however, seems to place causal primacy—at least in the case of an agent manifesting a non-causal power that results in change—outside of the realm of substances and into the realm of relations. It is now no longer the agent herself causing change, that is, the agent substance’s *intrinsic* causal power. It is the agent’s *extrinsic* non-causal manifestations triggering the manifestation of other powers accidentally that are the real prime movers in any given causal scenario. This calls into question if whether causal effects can really *deterministically originate* within human agents.

A further conceptual problem also arises because, as a concept that is proposed to account for how an agent possesses incompatibilist freedom, non-causal power alienates the agent from her actions to an unacceptable degree. This alienation results in problems explaining (1) how an agent ensures a causal link between his willing and action, and (2) how any action can be rightly attributed to the agent who caused it. (1) is indicative of an unacceptable lack of causal power possessed by an agent. We want to be able to say that we as rational agents have a higher degree of causal ability than merely causing the outcomes we wish to achieve *per accidens*; we want to say we *determine* our actions. (2) has to do with action ownership, that is, *responsibility*. How can I be responsible and know that I am

---

<sup>72</sup> Lowe, “Substance Causation,” 157-158; E. J. Lowe, “On the Individuation of Powers,” in *The Metaphysics of Powers: Their Grounding and their Manifestations*, ed. by Anna Marmodoro (New York: Routledge, 2010), 19.

responsible for the outcome of my actions if my actions are not entirely within my causal control? Obviously, a lot of moral and ethical considerations ride on this ontological issue. But barring any debate on what degree of control is necessary for moral responsibility and accountability, it is intuitively believed that human beings *are* responsible for their actions. Non-causal powers seem to weaken the basis for such accountability.

More importantly for formulating a theory of general causation, the notion of non-causal power, especially as observed in Lowe's natural examples, falters in the same fashion as Hume's regularity account of causation: it mistakes an epistemologically valuable, higher-order expression—one that is good for belief formation and justification—for what is actually occurring on a causally fundamental level of causal powers. Yes, the ball may be moving from one point in space to another, and yes it may not crush an object during its trip. However, we cannot ignore the very real causal powers at work in this instance of causation in order to formulate multi-applicable causal generalities concerning the movement of rolling objects, and how they may or may not crush objects, however epistemically valuable this may be for the discussion of such activities. Moving forward, it is clear that the notion of non-causal powers should be dispensed with. But what about the notions of passive and active powers? How do we make sense of powers in relation to each other, or the entities that possess them? The next section seeks to ontologically define and categorize powers, in order to clarify the nature of causation in general.

#### **1.4 Making Sense of an Ontology of Powers**

In order to develop a reasonable ontology of powers, it is important to clarify what powers are. There seems to be a consensus amongst those who hold powers views of causation (or *disposition* views, depending on the philosopher in question)<sup>73</sup> that powers are

---

<sup>73</sup> I favor the “powers” terminology due to clarity it provides in highlighting the existence of certain properties possessed by substances, although I use both terms interchangeably. See note 2 of this chapter.

properties possessed by concrete particulars. There is not as much agreement, however, concerning whether or not all properties are powerful in nature, that is, whether or not there are such things as purely categorical or purely qualitative properties. Further, as was the trouble with Lowe's theory, questions surround how powers manifest themselves, whether or not there are causal pairings, so to speak, between different types of powers, or if these pairings are indeed subject to a Humean law-like regularity themselves. And last, questions surround what these concrete particulars that possess these properties are, fundamentally speaking—whether they are merely material objects or, as Lowe held, substances. Obviously, this last question is of significant importance to developing an adequate neo-Aristotelian view of substance causation, which is the goal here.

#### ***1.4.1 Powers and Qualitative Properties***

Consider a ball with the property of *sphericity*. Arguably, this property not only gives the ball a certain look to it—it is round in shape—but it also conveys to the ball the power to roll down an incline, such as in the previous thought experiment. It seems then that the one and same property gives the ball two features, one qualitative (i.e., a round shape) and one dispositional (i.e., the power to roll). Does this mean that, fundamentally speaking, all qualitative properties (i.e., categorical properties) are dispositional properties, and vice versa? Or are some properties purely qualitative?

Some theorists favor a qualitative or categorical interpretation of properties in part because dispositional properties are empirically unobservable. Armstrong, in arguing against the powers view of causation, insists that all properties, regardless of the seeming behavior they confer on their objects, are fundamentally categorical in nature.<sup>74</sup> Thus, he denies the metaphysical reality of powers. Similarly, Kristina Engelhard denies the existence of wholly

---

<sup>74</sup> D. M. Armstrong, *Sketch for a Systematic Metaphysics* (Oxford: Oxford University Press, 2010), 48-50.

dispositional properties, yet nevertheless holds that all properties have both categorical and dispositional features to them.<sup>75</sup> So unlike Armstrong, she maintains the reality of powers, but argues that the knowledge of a property's dispositional features can only be abstracted "from the fact or state of affairs which [the property] is part of."<sup>76</sup> John Heil also equates dispositional properties with qualitative properties, but he holds this relation as a strict identity relation.<sup>77</sup> He argues that the idea of objects without qualities makes no sense; qualities are how entities possess power (i.e., entities possess power through their qualitative properties).<sup>78</sup> Furthermore, without qualities, there are no individuators for powers, thereby precluding causal relations.<sup>79</sup>

Then there are those theorists who favor a dispositional interpretation of properties over a qualitative one. Alexander Bird argues all properties are indeed dispositional—*essentially* dispositional: the dispositional manifestation of a property is *de re* necessary for that property, determining its trans-world identity.<sup>80</sup> Lowe himself holds that qualitative properties are not what confer power, but nevertheless may be part of what grounds a power. Moreover, the relation between a grounding qualitative property and a grounded power isn't one-to-one.<sup>81</sup> To go back to an example of his discussed earlier, a ball possesses the power to

---

<sup>75</sup> Kristina Engelhard, "Categories and the Ontology of Powers: A Vindication of the Identity Theory of Properties," in *The Metaphysics of Powers: Their Grounding and their Manifestations*, ed. Anna Marmodoro (New York: Routledge, 2010), 41.

<sup>76</sup> *Ibid.*, 53. To consider the categorical aspect of that same property, the property should not be "abstract[ed] from the fact or state of affairs that it qualifies but take it as part of the fact in question involving one and the same property." See *Ibid.*, 53-54.

<sup>77</sup> John Heil, "Powerful Qualities," in *The Metaphysics of Powers: Their Grounding and their Manifestations*, ed. by Anna Marmodoro (New York: Routledge, 2010), 69-70.

<sup>78</sup> *Ibid.*, 60-63.

<sup>79</sup> *Ibid.*, 64.

<sup>80</sup> Alexander Bird, "The Regress of Pure Power?" *The Philosophical Quarterly* Volume 57, No. 229 (October 2007): 514-516.

<sup>81</sup> Lowe, "On the Individuation of Powers," 18-19.

roll because of its qualitative property *sphericity*. However, soap bubbles also possess *sphericity* but have no power to roll. Thus, *sphericity* itself cannot confer the power to roll; some other appropriate property is needed—such as *rigidness*—in order for an object to possess the power to roll.<sup>82</sup>

Whether (at least some) power properties find their grounds in qualitative properties or whether they can be considered independent properties in their own right does not seem to affect the view being developed here. What is important to an account of substance causation in general, however, is that the power is possessed and wielded intrinsically by the substance that possesses it. This is especially true since powers are an essential part of a substance's identity conditions.

#### ***1.4.2 The Manifestation of Powers***

How is it that the manifestation of a power results in the effect or effects that it does? For example, how is it that water's power to dissolve salt and salt's power to be dissolved in water actually ends up manifesting in the dissolution of salt in water? One way philosophers have accounted for this is through the notion of mutual manifestation.

The mutual manifestation thesis holds that each and every power or disposition is ontologically paired with its associated manifestation partner or partners and only by these dispositional pairings or groups do manifestations of power occur. Or, more simply, as Austin puts it, "there *is no such thing* as a manifestation of a *single* dispositional property."<sup>83</sup> A power can only manifest when it is manifested with its manifestation partner.

The upshot to holding to the mutual manifestation thesis, its adherents believe, is that the need to distinguish between true causes and background conditions in causal explanation

---

<sup>82</sup> Ibid.

<sup>83</sup> Christopher J. Austin, "Is Dispositional Causation Just Mutual Manifestation?" *Ratio* 29 (2016): 235-236.

is negated.<sup>84</sup> There is no limit to what can be considered a manifestation partner, and therefore all manifestations in any given causal scenario can be considered causally relevant.<sup>85</sup> Furthermore, appealing to mutual manifestation provides an account of how a single power can have a wide variety of effects—one power can have varying manifestations based on whichever manifestation partner is present.<sup>86</sup>

The notion of manifestation partners, however, is incompatible with a true substance view of causation. This is because the manifestation of power is not the activity of intrinsically located power properties. Instead, manifestation is solely an extrinsic matter. Toby Handfield admits as much when, describing the manifestation of a rock shattering glass, he states that “[t]he explanation for why the rock and the glass have these two mutually complementary powers is that there is *one* manifestation kind—the shattering of a rock by a glass—in which both *being a rock* and *being a glass* are constituent properties.”<sup>87</sup> It is clear by this view of causation that manifestation is not determined by anything intrinsic to the substances involved. Thus, according to the mutual manifestation thesis, it would seem that in a world without glass, the rock’s power to shatter glass is non-existent, since the rock’s power to shatter glass is constituted by properties of both the rock and the glass. Without the existence of glass, rock lacks this power.

---

<sup>84</sup> S.C. Gibb, “Mental Causation and Double Prevention,” in *Mental Causation and Ontology*, ed. by S.C. Gibb, E. J. Lowe, and R.D. Ingthorsson (Oxford: Oxford University Press, 2013), 198. See also Austin, “Is Dispositional Causation Just Mutual Manifestation?” 237.

<sup>85</sup> Austin, “Is Dispositional Causation Just Mutual Manifestation?” 237-238.

<sup>86</sup> *Ibid.*, 238. See also Gibb, “Mental Causation and Double Prevention,” 197.

<sup>87</sup> Toby Handfield, “Dispositions, Manifestations, and Causal Structure,” in *The Metaphysics of Powers: Their Grounding and their Manifestations*, ed. by Anna Marmodoro (New York: Routledge, 2010), 110.

The mutual manifestation thesis is also rife with conceptual problems. Austin, in his analysis, points out two: an epistemological one and a metaphysical one. Regarding the epistemological problem, Austin explains:

[I]f dispositional properties are to be individuated according to an entire array of wholly distinct possible manifestations, how could we ever come to know when we have an instance of a particular dispositional property? For we can no longer pick-out a disposition by means of knowing its characteristic manifestation conditions (nor its characteristic stimulus conditions, mind), because its “characteristic” manifestation is perhaps only its most frequent, or perhaps its most theoretically interesting manifestation.<sup>88</sup>

The epistemological problem may be worse than Austin is characterizing it, however, since we could never attribute to a power a characteristic manifestation: its manifestation could never be isolated from its partners, and we therefore would never know what functional/behavioral contribution it is making. The metaphysical problem is closely connected to the epistemological one: if a single power has so many qualitative ways it can manifest, what metaphysically ties these differing manifestations together in the single property they arise from? Proponents of the mutual manifestation thesis, Austin argues, have no explanation.<sup>89</sup>

Further, an ontology where everything is causally relevant leaves everything of equal causal importance. This is especially true if we take all manifestations in a causal scenario to be contributing to the outcome of the event in question. But do we really want this to be the case? It would seem to hinder our ability to provide useful causal explanations, and, further, it would leave our intuitions concerning what is useful about these explanations unaccounted for. For example, Maslen argues that causal explanations that are more explicit are more

---

<sup>88</sup> Austin, “Is Dispositional Causation Just Mutual Manifestation?” 240-241.

<sup>89</sup> *Ibid.*, 241-242.

pragmatic.<sup>90</sup> He advocates utilizing Yablo’s process of “screening” in order to arrive at more pragmatically valuable causal explanations. Take for instance a death that is explained by “being hit by a red bus” that can be better explained by “being hit a by bus;” the fact that the bus was red can be “screened off” considering that if the bus was of any other color, the death would have still occurred; this process of screening helps us arrive at a better causal explanation.<sup>91</sup>

Whether or not true explanations can be arrived at reliably through this screening process is beside the point. What is important to recognize is the intuitive appeal this process has. This intuition is completely mitigated, however, if every property in a scenario is equally causally contributing and/or the effects of singular properties are epistemologically inaccessible. We would have no way of arriving at useful causal explanations except to provide a detailed account of every property involved in an event, such as including the fact that the bus is red in the example above. Proponents of mutual manifestation would have to be committed to the bus’ redness being causally relevant since to deny that it is is to distinguish between true causes and background conditions, which the mutual manifestation thesis isn’t equipped to do.

### ***1.4.3 Arriving at Causal Understanding***

Obviously, what is noteworthy about a substance view of causation—as opposed to merely a generic powers view—is the fact that what is doing the causing is a *substance*, as opposed to object or property. Why is this significant, and how does it help account for causation?

---

<sup>90</sup> Cei Maslen, “Pragmatic Explanations of the Proportionality Constraint on Causation,” in *Making a Difference: Essays on the Philosophy of Causation*, ed. by Helen Beebe, Christopher Hitchcock, and Huw Price (Oxford: Oxford University Press, 2017) 64-65.

<sup>91</sup> *Ibid.*, 58-59.

Appealing to substances, for one, helps in causal explanation, due in part to way substances retain their identity through change. As we have seen Lowe argue above, substances have, unlike objects, ontological independence, and therefore do not depend on any objects for their identity. This ontological independence also allows them to persist through non-essential material changes.<sup>92</sup> When we formulate causal explanations on the level of substances, we are able to track better the effects entities involved in causation are responsible for.

Relatedly, it is important to note specifically that artifacts cannot be substances. According to Richard J. Connell, true substances are the independently existing substratum in which properties adhere. Therefore, substances can change some (i.e., non-essential) properties—like shape—and remain the same substance. Artifacts, on the other hand, do not withstand this kind of change well and further depend on substances for their own existence.<sup>93</sup> In other words, unlike substances, artifacts lack ontological independence.<sup>94</sup> Hence, they cannot be substances.

Taking artifacts to be non-substances like this aids in identifying causal powers at work in any given causal scenario. As a case in point, in arguing against powers having an immediate role in the causal behavior of macro objects, Bird gives an example of a steel ball and a wooden ball being exposed to fire. One ball has the property being composed of steel, the other has the property of being composed of wood. One catches fire and the other does not. Bird concludes that the constitutive properties of the balls are causally relevant to one catching fire and the other not, but that neither of these properties can be dispositional since

---

<sup>92</sup> Lowe, “Substance Causation,” 154-155.

<sup>93</sup> Richard J. Connell, *Substance and Modern Science* (Houston, TX: Center for Thomistic Studies, 1988), 10-13, 25-27.

<sup>94</sup> See also David S. Oderberg, *Real Essentialism* (New York, NY: Routledge, 2007), 167.

they are both constitutional of the balls.<sup>95</sup> A better account of this causal event, however, would be to attribute properties not to the balls as objects, but to the wood and steel as substances. Wood has the property of being flammable at a lower temperature than steel catches flame, while steel is flammable (and melts as opposed to burns) at a higher temperature than wood. Powers of flammability belong to the *constitutive materials* that make up the balls, however—not to the *balls*. The shape of the balls—*spherical*—is also a contingent property held by both the individual wood and steel particulars. In that sense, Bird is getting the property-possessor relation backwards. What is doing the causal work is not the balls *per se* but the pieces of wood and steel which both happen to be ball-shaped. Taking the *ball*, an artifact, itself to be a substance—i.e., a possessor of power—would hinder us from identifying true causal dispositions at work. Such was also the problem with Lowe’s explanation of the causal powers at work with the ball rolling down an incline.

Appealing to a neo-Aristotelian view of substances also allows us to take final causes into consideration, which is conducive to better causal understanding. What are final causes? David S. Oderberg, in defining final causes, distinguishes them from efficient causes by specifying that they “are the *precondition* of the very possibility of any efficient causality.”<sup>96</sup> He goes on to add:

If fire burns wood but not pure water, if beta particles can penetrate a sheet of paper but not a sheet of lead, this can only be because the agents are ordered to some effects rather than others: they each have their own finality, which restricts the range of their effects... Remove the finality and you remove efficient causation altogether.<sup>97</sup>

This ordering of effects was what theorists were getting at when they appealed to mutual manifestation. Mutual manifestation allowed for the idea that the quality of a manifestation

---

<sup>95</sup> Bird, “Overpowering,” 357.

<sup>96</sup> David S. Oderberg, “Finality Revived: Powers and Intentionality,” *Synthese* 194 (2017): 2396.

<sup>97</sup> *Ibid.*, 2396-2397.

was due to the presence of a certain pairing or group of powers. The mutual manifestation thesis, however, was inherently untenable, and further, unfit for the purposes of explaining substantial causation since it drew too much of what causally motivated an event outside the substances involved. The notion of final cause, however, does much to explain the ordering of effects and, further in its favor, appeals only to the essential nature of a substance in question, thereby keeping all causal motivation intrinsic.<sup>98</sup>

It seems therefore that with a notion such as final causation, causal pairings do not need to be appealed to at all. This is even counter to Lowe's view of "active" and "passive" causal power pairings. We can conclude that all that is necessary for an effect to occur is the triggering and manifestation of a power, which itself is contained wholly within a substance. The triggering conditions and the quality of the manifestation is determined wholly by the essential nature of the substance whose power it is. Granted, the triggering may occur as a response to the presence of powers belonging to one or more other outside substances. However, the manifestation of a power is wholly a product of the substance in question whose power it is, however mixed in with other manifestations it may be during a causal event. And thus, a reasonable theory of neo-Aristotelian substance causation can be provided, one that reflects a reasonable metaphysical account of causation while allowing for the generation of better causal explanations than any Humean regularity theory could generate.

---

<sup>98</sup> Much more will be said about final causes, especially regarding their role in a moral ontology, in chapter 6.

## 2. The Metaphysics of Human Agency

### 2.1 Introduction

In Chapter 1 I briefly discussed human agency in the broader context of defending my view of a powers ontology as well as a point of contention with E. J. Lowe's notion of noncausal power. I briefly mentioned in that chapter that, in order to preserve his theory of incompatibilist free will, Lowe rejects that beliefs and/or desires causally factor into rational action. Indeed, if a belief or desire of an agent determines her action, then the integrity of self-determination—and any ability to do otherwise in any given instance of decision-making—is called into question.<sup>1</sup> However, does not rejecting this role of internal states render agent action arbitrary? If agent actions are arbitrary, then they cannot be rational.

In this chapter, I will expand on the topic of human agency, developing an account of agency that allows for the agent to be both rational and free. I will reject reasons internalism on the same basis that Lowe does, in that it bars incompatibilist human freedom.<sup>2</sup> Instead, borrowing from Lowe, I will account for rational action via reasons externalism, based on Jonathan Dancy's notion of reasons as states of affairs. The goal of this chapter is to provide an account of human agency and rational action that allows for the agent to be completely self-determined, i.e., undetermined by causal forces—including desires—outside agential control.

I will start with an analysis of reasons internalism, beginning with Bernard Williams' seminal theory. Because this theory calls into question incompatibilist freedom, I will look at attempts to rectify problems, focusing on Timothy O'Connor's causal analysis of the role of desires. I will reject O'Connor's theory on the basis that it fails to reconcile freedom, both

---

<sup>1</sup> E. J. Lowe, *Personal Agency: The Metaphysics of Mind and Action* (Oxford: Oxford University Press, 2008), 175-176.

<sup>2</sup> Ibid.

compatibilist and incompatibilist, with an internalist understanding of rationality. Then I will provide an account of externalism, largely adopting Lowe's view of rational agency. I will also provide an account of states of affairs as motivating, utilizing Dancy's analysis.

Note that for the purposes of this chapter, I am concerned with the question of reasons as causes or causal conditions for action, and thus whether reasons are causally efficient mental states.<sup>3</sup> More specifically, I will be determining whether desires are the causal origin of human action, or, in other words, if motivating reasons are desires. So whereas the literature on agential reasons reflects a much broader philosophical issue, the concern with the topic of reasons here is narrowly focused on the metaphysical status of so-called "motivating" reasons and how to reconcile reasons with incompatibilist free will. Any other issue having to do with reasons, such as their normativity—e.g., their rightness or wrongness given a certain situation or agent, etc.<sup>4</sup>—will be secondary to the metaphysical issue at hand.

## 2.2 Reasons Internalism

According to Bernard Williams' theory of rational action, an agent has a reason to perform an action if the agent has some desire that would be satisfied by performing said action, or a belief that his desire would be satisfied by performing said action.<sup>5</sup> The way that deliberation serves practical reasoning, according to Williams, consists in either finding new actions which one has a reason to perform or in finding new reasons to perform a certain action under review.<sup>6</sup> It is through this internal conception of reasons that Williams attempts to account for both the causal and rational origin of agential effects.<sup>7</sup>

---

<sup>3</sup> Maria Alverez, *Kinds of Reasons: An Essay in the Philosophy of Action* (Oxford: Oxford University Press, 2010), 50.

<sup>4</sup> *Ibid.*, 9.

<sup>5</sup> Bernard Williams, "Internal and External Reasons," in *Moral Luck* (Cambridge: Cambridge University Press, 1981), 101.

<sup>6</sup> *Ibid.*, 104.

<sup>7</sup> *Ibid.*, 102-103.

Reasons internalism is in contrast to reasons externalism, which Williams understands to be the position of taking of reasons to be independent of agent desire.<sup>8</sup> According to Williams, externalism is an untenable position:

If something can be a reason for action, then it could be someone's reason for acting on a particular occasion, and it would then figure in an explanation of that action. Now no external reason statement could *by itself* offer an explanation of anyone's action... The whole point of external reason statements is that they can be true independently of the agent's motivations. But nothing can explain an agent's (intentional) actions except something that motivates him so to act. So something else is needed besides the truth of the external reason statement to explain action, some psychological link; and that psychological link would seem to be belief. *A's* believing an external reason statement about himself may help to explain his action [emphasis original].<sup>9</sup>

Williams goes on to argue that belief alone cannot motivate action.<sup>10</sup> Therefore, Williams maintains that an externalist account of reasons needs to explain how certain beliefs can beget new motivations if it is to be taken seriously. However, Williams does not think such an account is possible without having to utilize internalistic claims, and as such externalist accounts are liable to collapse into internalism.<sup>11</sup> He concludes that the only genuine reasons for action are internalistic ones,<sup>12</sup> that is, those that consist in desires.<sup>13</sup>

The difficulty with internalism, as presented without further development, is that the notion of reasons as desires—or beliefs about desires—is in natural tension with free, self-determining human agency. This tension arises from holding that rational actions are only

---

<sup>8</sup> Ibid., 101.

<sup>9</sup> Ibid., 106-107.

<sup>10</sup> Ibid., 107-108.

<sup>11</sup> Ibid., 108-109.

<sup>12</sup> Ibid., 111.

<sup>13</sup> Ibid., 101. See also Neil Sinclair, "Reasons Internalism and the Function of Normative Reasons," *Dialectica* 71 (2017): 209. Sinclair defines internalism as the position that affirms a necessary conceptual connection between a normative reason and an agent's having rational cognitive access to existing motivation. It is the position of this paper, however, that to be motivated is just to be in a desire state, as will be clarified below. As such, defining internalism as the position that holds that reasons are desire states, as Williams does, is a simple and accurate reflection of the internalist position.

those actions that serve to fulfill our desires. As Joseph Raz points out, in the simplest understanding of the desire-belief model of reasons, practical reasoning would have to consist solely in determining what one desires to do most.<sup>14</sup> In this sense, our strongest desires, in a primitive way, determine our actions. It seems like we would not even have the power to refrain, since refraining itself would have to be born of our desires to refrain. So unless my desire to refrain from performing is stronger than my desire to perform some action, it is certain I shall perform said action. Otherwise, I would be irrational since I would be acting contrary to my only reason for action. Due to this constraint on rational action, I do not have the ability to do otherwise let alone choose my course of action in the incompatibilist sense.<sup>15</sup> Thus, I lack incompatibilist free will.

Of course, many internalist theories aren't this simplistic. Michael Smith's theory, for example, holds that whether or not we should perform some action depends on what "our fully rational selves" would advise us to do given our actual situation.<sup>16</sup> Further, the goal of deliberation is not to determine which desires to drop or adopt *per se*, but to build a coherent set of desires.<sup>17</sup> The way this set is formed is by determining whether our desires are "systematically justifiable." As Smith explains:

What do I mean when I say that we sometimes deliberate by trying to find out whether our desires, as a whole, are systematically justifiable? I mean just that we can try to decide whether or not some particular underived desire that we have or might have is a desire to do something that is itself non-derivatively desirable, and that we do this in a certain characteristic way: namely, by trying to integrate the object of that desire into a more *coherent* and *unified* desiderative profile and evaluative outlook.<sup>18</sup>

---

<sup>14</sup> Joseph Raz, *Engaging Reason: On the Theory of Value and Action* (Oxford: Oxford University Press, 2002), 50.

<sup>15</sup> See Robert Kane, *A Contemporary Introduction to Free Will* (Oxford: Oxford University Press, 2005), 120-121, for a discussion of how incompatibilism requires both the ability to do otherwise and the ability to originate a course of action.

<sup>16</sup> Michael Smith, "Internal Reasons," *Philosophy and Phenomenological Research* 55 (1995): 112.

<sup>17</sup> *Ibid.*, 113-115.

<sup>18</sup> *Ibid.*, 114.

This coherence and unity of desires out of which the agent operates itself gives the set of desires rational preference; moreover, it guards against holding onto or adopting desires that cannot rationally fit within the desire set.<sup>19</sup>

Unlike Williams' theory, Smith's theory isn't agent relative. In other words, Smith doesn't hold along with Williams that individual desires, privately held, only have rational force for individual agents that possess them. On the contrary, Smith believes that the rational requirement for desires to be coherent in each individual agent rationally ensures, in the Kantian sense, that all desires of all agents will converge onto a single desire set; therefore all desires will have rational force for all agents.<sup>20</sup> Any given reason will be a reason for everyone, although reasons may apply differently to different people relative to given circumstances.<sup>21</sup>

If we understand internalism correctly, then we have to conclude that rational actions are those actions that fulfill our desires. Therefore, the notion of our fully rational selves would be the notion of ourselves as acting in order to perfectly serve our fully informed and maximally coherent desires. We still have not gotten away from the idea that our desires are so deterministic that we lack the ability to do otherwise and, by extension, lack free will. We still seem to be brutally following our desires, such as with Raz's simple version of the belief-desire model.<sup>22</sup> Smith almost gets around this criticism by appealing to a coherent desire set that arises via systematic justification. We would have a commitment to our desires so long as they are coherent, and acting rationally would be acting to fulfill the desires of our

---

<sup>19</sup> Ibid., 115.

<sup>20</sup> Ibid., 118-120.

<sup>21</sup> Ibid., 120-122.

<sup>22</sup> Raz, *Engaging Reason*, 50.

rationally formed desire set. However, we still wouldn't have the ability to do otherwise during a decision-making event in light of our desires.

A further problem is accounting for how systematic justification is able to override desire, which it would need to do in order to form a coherent desire set. Upon reflection, Smith's commitment to systemic justification, in order for it function, would end up running counter Williams' contention that belief itself has no ability to motivate.<sup>23</sup> Indeed, how can a belief about a desire, regarding its coherency with other desires, override said desire? It could only do so if it was self-motivating or accompanied by a desire to have coherent desires. If it was self-motivating, then there's no reason to be committed to systematic justification and the existence of a desire set to motivate an agent. And if it was accompanied by a desire to have a coherent desire set, then agents in this case still merely brutishly follow their desires. So what's to prevent agents from brutishly following their desires in all cases? Why appeal to systematic justification and the formation of a desire set at all? Moreover, given that desires are the only things that can motivate, how can it be assumed that all agents are converging on the same desire set given that it would need to be our overriding desire for coherent desires, not rationality, that dictate what desires we adopt? It seems as if Smith's theory may also falter at the point of desires being deterministic. More needs to be said about the causal role of desire in order to make sense of an internalist schema that retains some semblance of self-determining agency.

### ***2.2.1 O'Connor's Reasons Internalism***

Like Williams, O'Connor defines reasons as agential desire (however, he claims his theory can work with other conceptions of internal reasons).<sup>24</sup> He lays out the conditions for rational action (in terms of desire) as follows:

---

<sup>23</sup> Williams, "Internal and External Reasons," 107-108.

<sup>24</sup> Timothy O'Connor, *Persons and Causes: The Metaphysics of Free Will* (Oxford: Oxford University Press, 2002), 85-86.

The agent acted then in order to satisfy his antecedent desire that  $\Theta$  if

1. prior to this action, the agent had a desire that  $\Theta$  and believed that by so acting he would satisfy (or contribute to satisfying) that desire;
2. The agent's action was initiated (in part) by his own self-determining causal activity, the event component of which is the-coming-to-be-of-an-action-triggering-intention-to-so-act-here-and-now-to-satisfy- $\Theta$ ;
3. concurrent with this action, he continued to desire that  $\Theta$  and intended of this action that it satisfy (or contribute to satisfying) that desire; and
4. the concurrent intention was a direct causal consequence (intuitively, a continuation) of the action-triggering intention brought about by the agent, and it causally sustained the completion of the action.<sup>25</sup>

Condition (2) is what O'Connor considers to be the agential decision point;<sup>26</sup> this is where indeterminism supposedly comes to bear on agential action.

O'Connor clarifies how these conditions work in tandem by addressing the Davidsonian contention that a reason for an action held by an agent prior to performing said action can still fail to be the reason for which said action was performed. Through an exposition of his conditions, O'Connor explains how intentions are freely formed and how, once formed, they cannot fail to result in action:

The answer [to Davidson's challenge] draws on all of the conditions. First, note that condition (3) is met only if I intend of my action that it satisfy the desire in question. (Recall that the relevant kind of intention is precisely an intention that triggers and sustains the action.) If I am conscious of a desire that I believe would be satisfied by my action but do not intend that my action satisfy it—maybe it is not a great concern to me or I have a desire that conflicts with it—then I do not truly act to satisfy it. Conditions (2) and (4) ensure that the intention is not undercut or short-circuited in some way, thereby making it explanatorily irrelevant. In connection with (2), recall my argument in chapter 3 that an agent-causal event could not itself be directly produced, either by the agent himself or by any other factor. This precludes some external factor's *initiating* my action, independent of my having had those reasons to which the intention refers. Condition (4), by requiring that that intention causally sustains the completion of the action, suffices to rule out any scenario in which some external factor takes over at this later state, independent of my having begun to act on reasons I had prior to acting.<sup>27</sup>

---

<sup>25</sup> Ibid., 86.

<sup>26</sup> Ibid.

<sup>27</sup> Ibid., 87.

Thus, Davidson's challenge does not even come to bear in that an agent's intention, and thus reason, cannot fail to result in the action that the agent intends.

Another Davidsonian issue that O'Connor deals with consists in the claim that the agent can be wrong about her reasons especially when she has two or more in conflict. This would call into question an intention's ability sustain the action, as per condition (4), which would seemingly be mitigated if the agent was mistaken about her reasons for acting.

O'Connor addresses this issue by admitting that the agent's awareness of her motivations comes in degrees, and, further, that the agent sometimes is not aware of the reason or reasons for which she acts. Since she can be genuinely mistaken or even deceive herself about her reasons, she may not know the contents of her action-triggering intention. As a consequence, O'Connor concedes that if the agent is wrong about her intention, then she does not act freely.<sup>28</sup>

O'Connor goes on to deal with the case of the agent believing that her reasons did not favor her chosen action, but favored another course of action—yet, despite these reasons, the agent chose her action. This would seemingly force us to judge that the agent's action was arbitrarily chosen, an undesirable assessment when trying to give an account of action that is rational. In response to this, O'Connor argues that his theory does not, in fact, require that the agent's action is judged arbitrary in this circumstance. The restrictions on rational choice that would warrant such a judgment are too great; they would not leave room for the possibility of more than one equally rational choice is available to the agent.<sup>29</sup> O'Connor points out that when selecting a gallon of milk from a row of indistinguishable gallons of milk at the supermarket, the choice is seemingly arbitrary: no one gallon of milk is better than the

---

<sup>28</sup> Ibid., 87-88.

<sup>29</sup> Ibid., 89-90.

other.<sup>30</sup> It would be wrong, however, to say that the picking of the one chosen gallon of milk over all the others renders the whole action of picking a gallon of milk arbitrary; the agent still performs the action for a reason since the agent desires milk, and the action is considered rational.

Relatedly, O'Connor also addresses the common conflicts that arise between moral duties and desires. The example he gives is that of deliberating between working or spending time with a friend in need. Perhaps one is too absorbed with work and should therefore spend more time with friends, etc. However, O'Connor argues that whatever one chooses in this instance would be "subjectively rational" since the action "fits" with one's character.<sup>31</sup> O'Connor further extends this problem to include when the agent is confronted with the possibility of two or more courses of action—each with different *kinds* of motivations but of the same level of reasonableness. The example he uses is choosing between what leisure activities to engage in while taking a break from work. Again, O'Connor argues that whatever one chooses, since it is in line with one's character, is rational.<sup>32</sup> A subtype of this case, according to O'Connor, is when the agent has many conflicting, yet equally regarded, values. It is in this instance that, in order to account for agent rationality in this type of decision-making, O'Connor admits the agent needing to rank values.<sup>33</sup>

To account for the importance of reasons in agent causation without giving them causally determining power (with the hopes of thus maintaining agential free will), O'Connor argues that reasons fulfill a structuring role, as opposed to a causally efficient role, in agent causation.<sup>34</sup> This structure also consists in persistent states of the agent's character, which

---

<sup>30</sup> Ibid., 90.

<sup>31</sup> Ibid.

<sup>32</sup> Ibid., 90.

<sup>33</sup> Ibid., 90-91.

<sup>34</sup> Ibid., 95.

include the agent's "relatively fixed dispositions and long-standing general intentions and purposes."<sup>35</sup> According to O'Connor, an agent's action arises out of and is guided by her structured propensity. Specifically, the agent exercises her free power when she generates her intentions on the basis of her propensity. O'Connor explains this as follows:

[I]t is as *cause* that the agent may be said to carry a propensity to *generate* a specific type of intention. We may suppose, that is, that recognizing a reason to act induces or elevates an objective propensity of the agent to initiate the behavior. One large independent advantage of thinking of tendencies as being carried in this way on the cause, not the effect, side of the equation, is that we needn't mysteriously invoke chance when asked to explain why a particular tendency is realized. Instead, the agent himself brings about the intention's obtaining.<sup>36</sup>

Since it is the agent herself bringing about the intention, and not the reason, the agent maintains free agency. As O'Connor clarifies further:

In any plausible construal of the causal theory of action, the agent's coming to have, or to recognize, reasons to act do not immediately trigger the intention. They set in motion a deliberative process that involves many other factors and culminates in forming the intention... The claim [being made here], in the position now entertained, is that when the wider *context* is right—when the agent is so internally constituted and externally situated as to be capable just then of exerting active power—then the uniformly *characteristic* influence of the state of having or recognizing reasons to A is ultimately transmitted into an influence on the agent's causal capacity, giving it a propensity structure in the manner indicated.<sup>37</sup>

Although it is not entirely clear what O'Connor means by "influence," it seems O'Connor is arguing that since reasons don't have a direct role in causation then they can be considered not strictly determining. What does have a direct causal role, according to O'Connor, is the agent's own "causal capacity," seemingly exercised through intentions that the agent forms. Thus, the agent remains a free self-determining agent.

---

<sup>35</sup> Ibid., 97-98.

<sup>36</sup> Ibid., 97.

<sup>37</sup> Ibid., 99.

O'Connor also considers the scenario where the agent has only one reason to commit one action, a scenario which would seem to result in the said reason being deterministic. O'Connor answers this problem by reference to Duns Scotus, who held that an agent always has the ability to refrain from an action. However, the question remains as to whether an agent has a reason to refrain, since if she does not, she would not. O'Connor simply replies that the agent could possibly refrain for a time.<sup>38</sup> With this answer, however, there still remains a question of whether or not this action is ultimately causally determined. In response, O'Connor admits that his view does not include "that agents always act freely, only that when they do, their acting was not then inevitable."<sup>39</sup> Thus, in a scenario where the agent has only one reason, even O'Connor himself admits that her freedom is largely limited.

### ***2.2.2 Problems with O'Connor's View and Internalism in General***

Problems with O'Connor's view stem from his basic project of trying to give an explanation of actions in terms of an internalistic conception of reasons while maintaining that reasons do not cause action. This would allow agents' actions to be explainable via internal states—and therefore rational—but also allow agents themselves to be causal originators of their actions—and thus free. Giving an account of how actions can causally originate with the agent and yet be explainable via internal states is the basic project of any internalist trying to maintain some semblance of agential freedom.

The way O'Connor addresses the issue at hand is through specifically analyzing the role of the agent's desire in the formation of an intention. Recall O'Connor's conditions for reasonable action:

The agent acted then in order to satisfy his antecedent desire that  $\Theta$  if

1. prior to this action, the agent had a desire that  $\Theta$  and believed that by so acting he would satisfy (or contribute to satisfying) that desire;

---

<sup>38</sup> Ibid., 100.

<sup>39</sup> Ibid., 101.

2. The agent's action was initiated (in part) by his own self-determining causal activity, the event component of which is the-coming-to-be-of-an-action-triggering-intention-to-so-act-here-and-now-to-satisfy- $\Theta$ ;
3. concurrent with this action, he continued to desire that  $\Theta$  and intended of this action that it satisfy (or contribute to satisfying) that desire; and
4. the concurrent intention was a direct causal consequence (intuitively, a continuation) of the action-triggering intention brought about by the agent, and it causally sustained the completion of the action.<sup>40</sup>

Conditions (1) and (2) seem to indicate that the agent's reasons merely consist in the desires that are the focus of an agent's intention formation. Condition (2) precludes anything, including the agent herself, of forming an intention independent of her desires—otherwise the subsequent action is considered unreasonable. In order to give an account of *free* agency, O'Connor claims, as per condition (2) that the agent generates her intention in virtue of the reasons she has, that is, in virtue of her desires. So, under O'Connor's theory, it is in the generation of intentions that both reasons come to bear and agential freedom comes into play.

According to O'Connor's view, it is irrational for an agent to form an intention counter to her reasons—reasons being her desires that factor into forming of intentions. More specifically, O'Connor's definition of free action, according to his conditions, requires that an agent act in accordance with her desires in order for her actions to be considered not only reasonable, but *self*-determined. Any action considered free and reasonable will involve intention-generating desires, making it impossible for an agent *not* to form an intention in accordance with her desire if we are to consider her “free” and “reasonable.” There are two ways we can preserve free action given this rational stipulation: 1) give the agent the ability to choose between desires, and 2) allow the agent to refrain from fulfilling her desires. O'Connor attempts both avenues but subsequently fails to preserve freedom. This failure is a failure of internalism generally.

---

<sup>40</sup> Ibid., 86.

Concerning the first avenue, the ability to choose between desires: it would only be reasonable, in the sense that O'Connor theory explicates reason (as desire-based), to allow strong desires to negate weaker ones. If we didn't allow this, we would be basing rationality of action on something other than desire, thereby negating O'Connor's whole project. Therefore, the agent would *always* be rationally obligated to act in accordance with her strongest desire at any given moment. If she fails to act on her strongest desire, her action cannot be considered reasonable, according to O'Connor's internalist definition of reasonable. And if the agent's action cannot be considered reasonable, it cannot be considered free.

Moreover, according to O'Connor's conditions (2) and (4), the action-triggering intention—which is caused by the agent—causes a concurrent intention necessary to follow said action to completion; consequently, the agent has no power to put a stop to an intention once it is formed. Thus, even if we were to accept it is the case that a desire does not lead to action is when a stronger desire overrides it (and that this overriding is sufficient for agential freedom), the agent needs to be aware of her stronger desire before she forms her intention. Otherwise, if her intention is formed before full awareness of desires is reached, the agent is mistaken about her intention. And O'Connor does not allow the agent to be mistaken about her intention and retain agential freedom.<sup>41</sup> Therefore, our actions must *always* be determined by our strongest desires in order to be considered free—yet, this is incompatible with the notion of freedom we're trying to preserve. It seems as if O'Connor's view collapses into Raz's crude and simplistic version of internalism, an undesirable outcome.

O'Connor might counter that the agent can still be free to choose between competing desires of equal strength. However, this seems to confine freedom, according to O'Connor's

---

<sup>41</sup> Ibid., 88.

view, exclusively to the domain of cases like taking milk off the dairy case in the supermarket, that is, to cases of choosing between completely indistinguishable low-value options. This is because, arguably, the set of desires surrounding anything more significant seems to have more contrastive gradations. Even if we did allow for the agent to choose between equally competing desires (thereby narrowing the scope of O'Connor's account severely), according to O'Connor conditions, the subsequent action would fail to be reasonable and thus free. This is due to the causal activity of an agent inherently *stemming from her reasons*, i.e., the commitment that internalists have to internal reasons as a necessary condition for causal activity. In the case of O'Connor, whatever would be overriding a competing desire *could only be another desire*, and not the agent herself, since a desire is needed for the activity to get off the ground, so to speak. Therefore, the agent would have only the option of following her strongest desires, and we've reverted back to Raz's crude notion of internalism.

Now concerning the second avenue, the agent's ability to refrain: in addressing the scenario where the agent only has one reasonable option available, O'Connor admits that all he strictly contends is that any action at any given moment is not inevitable.<sup>42</sup> In other words, O'Connor contends that his theory, *at the very least*, preserves the agent's ability to refrain from acting. Counter to O'Connor's contention, however, his theory does not even allow refraining. Since refraining would essentially be the agent refusing to act on her desire *without reason* (i.e., with no desire), it would be considered unreasonable—and subsequently unfree—on O'Connor's theory of free and rational action. The only way an agent could choose to refrain is if her desire to refrain was stronger than her desire to commit the action

---

<sup>42</sup> Ibid., 100-101.

in question, and we are once again back to Raz's undesirable caricature of internalism where the strongest desire wins.

Now it has become clear why internalism has a natural tension within it that is impossible to escape. In order for an action to be considered free, it must not be determined by internal states. In order for an action to be considered free, it must not be arbitrary, which would be inconsistent with its being rational. In order for an action to be considered rational, it must be determined by internal states. Thus, when internalist set out to demonstrate free will, they set out to demonstrate a contradiction: that actions can both be determined and not determined by internal states.

### **2.3 Reasons Externalism**

It is the very notion of rational actions as needing to originate from belief-desire states that leads to the above metaphysical problem; the internalist constraint on what can be considered rational is too restrictive, thereby leading to desire determinism. Either the agent sacrifices her freedom in order to be rational, or her rationality in order to be free. Instead of putting the agent in this bind, I argue along the same lines as Lowe when he states that "although a belief can sometimes cause or help to cause some of an agent's behavior, a rational action is precisely one that *isn't* caused by an agent's beliefs and desires. Rather, it is one that the agent *chooses* to perform 'in the light of' his or her beliefs and desires, which is a very different matter."<sup>43</sup> Therefore, a better conception of reasons is that of Lowe's, who holds that reasons are the external circumstances that bear upon the agent's consideration, giving an agent a certain awareness upon which she bases her judgments.<sup>44</sup> Because reasons are placed outside of the agent, the notion that a rationally disruptive break can occur

---

<sup>43</sup> Lowe, *Personal Agency*, 181.

<sup>44</sup> E. J. Lowe, "The Will as a Rational Free Power," in *Powers and Capacities in Philosophy: The New Aristotelianism*, ed. by Ruth Groff and John Greco (New York: Routledge, 2013), 181.

between an agent and her reasons—which was seen consistently demonstrated in the above analysis of O’Connor’s theory—is mitigated. Externalism allows the agent’s actions to be considered more consistently rational and the agent herself more reliably free.

According to Lowe’s theory, which largely makes use of Jonathan Dancy’s work,<sup>45</sup> what makes an action rational is that the action is performed *in light of* a reason that was *chosen* by the agent.<sup>46</sup> This choosing of a reason is a freely made choice by the agent—a choice uncaused by prior events and able to causally determine future events.<sup>47</sup> Hence, on Lowe’s theory, reasons themselves do not risk becoming causally determining; the agent herself is self-determining in choosing the reasons in light of which she acts. This, importantly, maintains incompatibilist freedom under externalism.<sup>48</sup> Moreover, the reason “is a fact or state of affairs which justifies a particular course of action,”<sup>49</sup> and hence is rationally able to explain said action.

This view does not deny that beliefs and desires can have any role in agential deliberation. It merely places the criterion of what makes an action rational outside the internal states of the agent. As Lowe explains:

[I]t is not the fact that *S* believes that the slate is falling [off the roof] that constitutes a reason for *S* to jump out of the way: rather, it is the fact that the slate *is* falling—although, once more, *S* must be *aware* of this fact and so believe, or at least suspect, that the slate is falling if he is to be in a position to act upon that reason. The mistake that is characteristically made by those who favour an ‘internalist’ view of reasons for action is to regard such psychological states *themselves*, rather than the *contents* of such states, as constituting reasons for the agent to act—where those ‘contents’ are the *facts or states of affairs* whose existence is revealed to the agent in virtue of his being in such psychological states.<sup>50</sup>

---

<sup>45</sup> Lowe, *Personal Agency*, 181, 186, 199.

<sup>46</sup> *Ibid.*, 182.

<sup>47</sup> *Ibid.*, 183-184.

<sup>48</sup> *Ibid.*, 184.

<sup>49</sup> *Ibid.*

<sup>50</sup> *Ibid.*, 181.

So it isn't the psychological states of the agent themselves that constitute an agent's reasons, as internalism would hold, but the states of affairs that those psychological states give the agent access to.<sup>51</sup> The external circumstances that the agent finds herself in is what constitutes the reasons for which she acts in the way she does when she acts in light of them. As Maria Alvarez points out, internalists about the belief-desire model make the mistake of thinking that motivation needs to come from an internal state of desiring as opposed to the external things which are desired.<sup>52</sup> Internal states do not need to be invoked to account for an agent acting in light of a desire or belief.

What if the agent has false beliefs about her external circumstances and acts in light of this false belief? This case would seem to render any subsequent actions performed by the agent irrational. Lowe argues that the agent, in order to be rational, only needs to be aware of her reason for acting. However, this awareness "doesn't necessarily require that the reason should *actually* obtain."<sup>53</sup> As Lowe explains:

In [the case of the agent acting on a false belief], the agent might well have done better to have acted differently, but he did not act *irrationally*, since he still acted for a reason that he chose to act upon. Since reasons are not causes, it does not matter that an agent's reason for acting in a certain way may on occasion be a state of affairs which does not obtain and hence which is not a fact.<sup>54</sup>

In the case of a false belief, Lowe contends it could be argued that the agent acted for no reason, since a reason did not actually obtain; however, the agent would still have acted rationally.<sup>55</sup> Lowe justifies this position as follows:

We could say that if an agent jumps out of the way in the light of a *mistaken* belief that a slate is about to fall on his head, then he does not *really* have a reason to act in

---

<sup>51</sup> Ibid.

<sup>52</sup> Alvarez, *Kinds of Reasons*, 118-119.

<sup>53</sup> Lowe, *Personal Agency*, 184.

<sup>54</sup> Ibid., 185.

<sup>55</sup> Ibid., 185-186.

that way on that occasion. At the same time, however, we can and should insist that such an agent is not acting *irrationally* merely insofar as his belief is false—for if it had been true, he really *would have had* a reason to act in precisely the way that he does. On this view of the matter, only *facts*—that is, states of affairs which obtain—can actually *be* reasons for action, but an agent can, nonetheless, act for a *would-be* reason which does not actually obtain and in doing so act rationally in precisely the same sense in which an unmistaken agent can.<sup>56</sup>

Hence, Lowe makes a distinction between reasons, which are external states of affairs, and rationality, which is the agent's ability to apply reasoning skills to her perceived situation and determine her actions.<sup>57</sup>

Utilizing this external notion of reasons helps us arrive at better explanations of agent actions when those agents have diminished mental capability. For example, someone undergoing an episode of psychosis commits a murder. Perhaps she believed the victim was a reptoid (i.e., lizard person) in human disguise. The internalist would have to say that this event either occurred for no reason—since the agent's internal state was irrational, given her irrational belief—or that the reason it occurred was that the agent desired to rid the world of reptoids. The externalist, however, can attribute the murder to the murderer's ill mental health, which is an external factor more relevant to the explanation of her actions than her internal mental state.<sup>58</sup>

Someone like O'Connor may counter this by saying that, on his view, the agent has no discernible *rational* intention and therefore has acted involuntarily (rationality being a necessary condition for free action). This seems to cohere with how we think about actions born out of mental illness. However, O'Connor's view does not allow for mental fitness to be taken into consideration when explaining agent actions. For explanation of action, he only appeals to agent intentions, whether or not those are rationally aligned with the agent's

---

<sup>56</sup> Ibid., 185.

<sup>57</sup> Ibid., 182-183, 185. See also Alvarez, *Kinds of Reasons*, 24, 140-144. Alvarez also argues that the agent who acts in light of a false belief acts for no real reason, but for what she calls an "apparent reason."

<sup>58</sup> See also Lowe, *Personal Agency*, 180-182.

desires (and arguably the agent's intentions are rationally aligned with her desires to rid the world of reptoids). The most O'Connor could hope to do to account for mental illness is to state that it is part of the agent's propensity to behave a certain way. However, mental illness is not a propensity of an agent, especially if the illness is caused by external factors—such as chemical imbalances—that are unmitigated by agent habituation (habituation being part of what seems to establish an agent's intention for O'Connor).<sup>59</sup> A full and thorough explanation of the agent's action in this scenario cannot be given using O'Connor's internalism.

This highlights an important distinction to be made between reasons and causal explanation, as also pointed out by Alvarez. Those who hold to the internalist belief-desire understanding of reasons as causally accounting for agent action are conflating reasons with causal explanation. However, reasons and causal explanation are two different things in the same fashion that metaphysical causal accounts and the explanation of events are two different things (see chapter 1, section 1.2.3). As Alvarez argues, if one has a paranoid delusion that one is being followed by MI5, one has the apparent reason to go around in disguise. However, what one really has reason to do is see a psychiatrist. Nevertheless, that one has such a paranoid delusion helps explain one's behavior.<sup>60</sup>

In addition to providing a better account of action, reasons externalism is in better accord with our phenomenology as deliberators and agents. As Lowe contends, the phenomenal experience of rational agents doesn't seem to support the notion that beliefs and desires are the causes of our actions; it seems to us that we've freely chosen what to do in the face of our beliefs and desires.<sup>61</sup> This accords well with an experimental philosophy study—

---

<sup>59</sup> O'Connor, *Persons and Causes*, 97-98, 105.

<sup>60</sup> Alvarez, *Kinds of Reasons*, 29, 48-49.

<sup>61</sup> Lowe, *Personal Agency*, 188-189.

conducted by Oisin Deery, Matt Bedke, and Shaun Nichols—which was able to show that, through educating people on what philosophers mean by determinism and by utilizing their self-reports regarding deliberating between choices, people’s experiences of deliberation not only do not reflect determinism, but they also do not reflect compatibilism. Phenomenology of deliberation, according to Deery et al., is strictly an experience of incompatibilist deliberation,<sup>62</sup> even when deliberation involves choices that are moral in nature.<sup>63</sup> More pointedly, people reported their experiences of agency as reflective of being able to make decisions counter to their desires.<sup>64</sup> Phenomenology of deliberation thus supports an externalist, incompatibilist view of agency and not an internalist or compatibilist view.

What about the objection that, if our actions are free in the incompatibilist sense, and if our desires or beliefs do not have causal influence on our choices, then our choices are arbitrary? Lowe specifically addresses Peter van Inwagen’s thought experiment of an agent, Alice, who is able to go back and “replay” an instance of choosing between two available choices—one to tell the truth, the other to lie. Opponents of incompatibilism will argue that the probability of any one of the agent’s choices coming up again in the replay runs from unity (strictly determined) to 50/50 (pure chance).<sup>65</sup> What is clear, they argue, is that the agent herself does not have determining ability over her choices—she is either determined from without or her choice is arbitrary. Lowe answers this objection as follows:

The way in which the thought experiment is set up requires us to suppose that in all ‘replays’ of her situation her deliberations proceed in exactly the same way that they actually did, but that at the end of some of these replays she still chooses to tell the truth while at the end of others she chooses instead to tell a lie. However, if Alice deliberates and chooses rationally, at the end of her deliberations she will have

---

<sup>62</sup> Oisin Deery, Matt Bedke, and Shaun Nichols, “Phenomenal Abilities: Incompatibilism and the Experience of Agency,” in *Oxford Studies in Agency and Responsibility, Volume 1*, ed. David Shoemaker (Oxford: Oxford University Press, 2013), 148.

<sup>63</sup> *Ibid.*, 139.

<sup>64</sup> *Ibid.*, 145.

<sup>65</sup> Lowe, *Personal Agency*, 190-191.

formed a judgement as to whether the reasons in favour of telling the truth are better than those in favour of telling a lie. Now, certainly, if Alice's choice was genuinely undetermined, it was possible for her to have chosen to tell a lie in spite of having formed such a judgement. On the other hand, there is a clear sense in which, in the light of that judgement, Alice would have been acting irrationally in telling a lie. A libertarian should happily accept that our freedom to choose is a freedom even to choose irrationally—and we all recognize, I think, that sometimes we do make irrational choices.<sup>66</sup>

I do not agree with Lowe at this point. Conceding that the agent makes an irrational choice in this instance falls to the objection that, if the agent fails to make a rational choice, then the agent fails to be free (rationality being a necessary condition for agential freedom). As such, I think Lowe is assuming a narrower view of rational action than is necessary.

Counter to Lowe, I argue that the choice made by the agent, no matter what the choice is, is rational as long as the choice is made in light of some reason.<sup>67</sup> In order to maintain this position, I argue that the rationality of an action admits of degrees and, as such, some reasons for action are better suited to a situation than others. Take for instance an individual who is deciding whether to focus on athletics instead of academics in high school. This agent knows that being a professional athlete pays more than a white-collar job, and therefore she has a reason to focus on athletics. However, choosing to focus on academics is a safer bet since landing a white-collar job is significantly easier than getting a sports contract. It would be tempting to say that focusing on academics is the rational choice while focusing on athletics is irrational, given all available information. Nevertheless, her actions, no matter what she chooses to do, are rational since there is some reason for each course of action. Hence, even

---

<sup>66</sup> Ibid., 192-193.

<sup>67</sup> Those who are concerned with normative reasons may have issue with my definition of rationality here given that they may hold *bad* reasons are *irrational* reasons. I respond that some bad reasons border on irrationality; however, I want to make a distinction between a completely irrational reason—and thus a non-reason leading to non-free action—and a bad reason. For example, someone whose reason for getting a haircut is because his favorite color is blue is acting irrationally. On the other hand, someone whose reason for getting a haircut is because he wants to make all his co-workers envious is acting for a bad reason. For one, it is arguably morally not good to try to make others envious. Secondly, it is probably the case that none of his co-workers care about how he looks anyways and therefore they will not be prone to envy. Still, this agent is acting rationally—if it were the case that his co-workers were prone to envy, then perhaps an astounding haircut would cause them to be envious.

if the agent comes to believe that the more rational course of action is focusing on academics, if she chooses to focus on athletics she does not act irrationally.

This seems to cohere with how we make distinctions between agents and evaluate their actions, especially in criminal courts. There is a difference between completely irrational agents and agents who fail to act for reasons the law requires they act on. The former are not legally responsible for their actions while the latter are. In general, it seems that agents—who are more often than not rational—can always choose to act on reasons that are more or less proper to their circumstance (that is, the conditions for an action to be rational is easier to meet than is the conditions for said action to be considered proper). In this sense, even though agents usually always act in light of reasons, the quality of the reason they choose to act upon, or the quality of the course of action they choose in light of a reason, can vary. This explains why some courses of action are favored over others, but why more than one course of action may be considered rational given a situation.

#### **2.4 Motivating States of Affairs**

I now want to provide an account of states of affairs as motivating. To do so, I will draw upon the work of Jonathan Dancy. Dancy specifically addresses the problem of whether an externalist conception of reasons can be thought of as motivating in the same way that internalists take desires, fulfilling the role of reasons, to be motivating, considering that desires are the only kind of thing that seem to naturally motivate an agent.<sup>68</sup> Dancy seeks to provide an account of how reasons, as distinct from desires, can indeed motivate an agent to action and, furthermore, how reasons can only be thought of as states of affairs.

First, Dancy rejects the internalist assertion that, since only desires give rise to motivation, reasons must consist in desires. He does this by drawing out the distinction

---

<sup>68</sup> Jonathan Dancy, *Practical Reality* (Oxford: Oxford University Press, 2000), 10-12.

between desires and reasons. For starters, Dancy rejects the assumption that desires give rise to motivation to begin with,<sup>69</sup> clarifying that “to desire is just to be motivated.”<sup>70</sup> Then Dancy argues that desires, i.e., motivations, are formed in light of reasons, not reasons in light of desires.<sup>71</sup> As he points out, “if someone wants something that there is no reason to want, his desire does not give him some reasons for doing it, a reason that was not there before,”<sup>72</sup> and moreover, that “[i]f an action is silly, wanting to do it does not make it any less silly.”<sup>73</sup> Since desires don’t seem to determine the quality of reasons in the way that one would expect, but that reasons seems to dictate the quality of desires, we must conclude that reasons are distinct from desires, and further that reasons are the basis of desires.

Dancy concedes that there are inclination and urges, special kinds of desire that don’t seem to be based on reasons but are still the basis for some actions. However, he argues that inclinations only come to bear on agency when reasons fail to be decisive enough to determine action—moreover, urges without reasons are pathological.<sup>74</sup> This assessment seems to cohere with how we evaluate an agent who acts on impulse but who lacks any coherent reason, or even possesses a counter reason, for her action. Hence, neither inclinations nor urges make good candidates as reasons when giving an account of *rational* agency. Thus, Dancy is able to make a distinction between reasons and desires, one that is incompatible with internalism.

---

<sup>69</sup> Ibid., 13-14.

<sup>70</sup> Ibid., 14. See also Alvarez, *Kinds of Reasons*, 119-120. Alvarez clarifies that the internal state of desiring just is the internal state of being motivated, as opposed to desiring being an internal state that is motivating and thus separate from desiring.

<sup>71</sup> Dancy, *Practical Reality*, 35-36.

<sup>72</sup> Ibid., 36-37.

<sup>73</sup> Ibid., 37.

<sup>74</sup> Ibid., 36.

In order to argue his case that reasons, without the need of desires, are motivating, Dancy borrows the distinction that some philosophers make between motivating reasons and normative reasons. Dancy defines motivating reasons as the “considerations in light of which [the agent] acted...which we can think of as what persuaded him to [act].”<sup>75</sup> This seems to be the typical notion of reasons that we have been working with up to this point, that is, the expressed criterion of rational action: an agent acting in response to a reason. On the other hand, normative reasons provide normative explanations for actions. As Dancy explains, some reasons are normatively good reasons for action (i.e., they “speak in favour of the action,” in Dancy’s words) and some good reasons are better than others. These are normative reasons.<sup>76</sup> I claimed at the end of the previous section that actions can be more or less reasonable considering a given circumstance; this seems to reflect what Dancy says here about normative reasons. Along with Dancy, I deny that there is a real distinction between motivating and normative reasons, as my proposed theory in the previous section reflects. However, Dancy contends that this distinction must be utilized in order to demonstrate its ultimate unintelligibility.<sup>77</sup>

When we speak of normative reasons for an action, we speak of reasons that favor the performance of that action, i.e., external normative states. On the other hand, when we speak of motivating reasons, we speak of our beliefs or desires, i.e., psychological states, that gave rise to said action.<sup>78</sup> This presents a problem when giving an account of reasons by putting “a constraint on any theory about the relation between normative and motivating reasons,” that is, any theory must “show that and how any normative reason is capable of contributing to

---

<sup>75</sup> Ibid., 1.

<sup>76</sup> Ibid.

<sup>77</sup> Ibid., 99.

<sup>78</sup> Ibid., 100-101.

the explanation of an action that is done for that reason.”<sup>79</sup> Dancy calls this the *explanatory constraint*. An easy way of meeting the constraint, according to Dancy, is to hold that normative reasons explain the motivating reasons that explain the action. For example, an agent’s actions can be explained by his belief, and his belief itself can be explained by the truth of the belief. Thus, there is transitivity to explanations of agential action, what Dancy calls a “three-part story.”<sup>80</sup>

However, as Dancy points out, there is also the option of bypassing motivating reasons altogether. In order to do this, normative reasons would themselves have to be motivating. Dancy argues that this is indeed the case: the reason why an agent should take a certain course of action can also be the reason why he does. This reason will both motivate the agent and make the action the correct course of action to take. Furthermore, Dancy takes these reasons to be constituted by what the agent believes about a situation, i.e., its state of affairs. Thus, if states of affairs themselves can be demonstrated to be motivating, then belief-desire states do not need to factor into explanations of agent action, and externalism will be demonstrated to be a plausible alternative to internalism in this regard.<sup>81</sup>

Dancy presents three objections to the three-part story in favor of his bypass theory. In order to support his first objection, he proposes another constraint besides the explanatory constraint, what he calls the *normative constraint*. Dancy defines this constraint as “requir[ing] that a motivating reason, that in the light of which one acts, must be the sort of thing that is capable of being among the reasons in favour of so acting; it must, in this sense, be possible to act for a good reason.”<sup>82</sup> In other words, there must exist motivating reasons

---

<sup>79</sup> Ibid., 101.

<sup>80</sup> Ibid.

<sup>81</sup> Ibid.

<sup>82</sup> Ibid., 103.

amongst normative reasons. The reason why Dancy thinks this can be is because, if normative reasons are such that they can contribute to motivating reasons and thus help in explanation, then motivating reasons must be such that they can possess normativity. As Dancy argues, this new constraint makes the three-part story untenable in that the three-part story bars normative reasons from being amongst the motivating reasons that an agent can choose from.<sup>83</sup> Dancy doesn't deny that some psychological states have normative status and therefore can add to the normative evaluations of an agent's actions;<sup>84</sup> he merely denies that psychological states alone can carry normativity and thus normative evaluation cannot be done on the basis of psychological states alone.<sup>85</sup> Normative reasons and motivating reasons are just too ontologically distinct.<sup>86</sup>

Dancy's second objection is similar to the first. He points out our tendency as agents to explain our actions via good reasons, i.e., reasons that favor our acting. Again, Dancy highlights the fact that states of affairs, and not psychological states, are the kind of things that can favor action.<sup>87</sup> Further, Dancy clarifies that this isn't so much a new objection as it is a generalization of the first: if a reason for acting "must be the right sort of thing to be a good reason,"<sup>88</sup> then it must be normative. Moreover, since only normative reasons can be more or less good, then they must comprise all reasons and therefore be motivating.<sup>89</sup>

---

<sup>83</sup> Ibid., 103-104.

<sup>84</sup> Ibid., 104.

<sup>85</sup> Ibid., 105.

<sup>86</sup> Ibid., 106.

<sup>87</sup> Ibid.

<sup>88</sup> Ibid., 107.

<sup>89</sup> Ibid.

Dancy's third objection, borrowed from Arthur Collins, holds that an explanation of an action cannot possibly be given by an agent without her committing to the truth of the beliefs she counts as relevant to the explanation. This is because the role of an explanation is not to convey the agent's perspective *per se*, although explanations are formulated from the agent's perspective. It is rather that explanations seek to clarify the state of affairs that an action was performed in light of.<sup>90</sup> In order to argue for this claim, Dancy makes a distinction between an explanation and a *psychologized* restatement of said explanation. A straightforward explanation is of the form "I am performing this action because *p*," while a psychologized restatement of the same explanation is of the form "I am performing this action because I believe that *p*." Dancy maintains that regardless of whether the agent offers the above straightforward explanation or the psychologized restatement, she is affirming the truth of *p*.<sup>91</sup> Dancy therefore argues that a psychologizing restatement of an explanation meets two conditions: 1) "[i]t does not introduce new subject-matter," i.e., a new thing to be explained, and 2) "[i]t does not delete the agent's endorsement," i.e., the ontological commitments expressed by the original statement.<sup>92</sup>

Under a three-part account of reasons and actions, however, a psychologized restatement made by a third-party cannot meet these two conditions, thereby rendering third-party explanations misleading. This is due to a disconnect of commitments between an agent's explanation of her actions and a third-party explanation of those same actions. Consider the nature of the parts of an explanation, according to the three-part theory: there is the factor that consists in the psychological state which gave rise to the action—what Dancy calls the *proximal* explanation—and the factor that consists in the circumstances that gave

---

<sup>90</sup> Ibid., 108.

<sup>91</sup> Ibid.

<sup>92</sup> Ibid., 109.

rise to the psychological state—the *distal* explanation. The three-part account attempts to leave third-parties free to provide proximal explanations without committing to the truth of any distal explanations. However, the agent herself cannot do this; whenever she provides an explanation, whether proximal or distal, the explanation assumes the truth of the beliefs expressed in the distal explanation. Thus, a third party psychologized restatement of an explanation will possess a different subject matter than an agent's own, rendering said explanation misleading if considered identical to the subject's own psychologized explanation.<sup>93</sup>

Dancy addresses the notion that third-party explainer should not take distal content into account when formulating explanations especially since an agent can be wrong about what she takes to be the case. Dancy answers that, regardless of the truth of distal content, to ignore it is to mischaracterize the agent's reason for her actions. In order to faithfully preserve the agent's reasons when formulating an explanation, we must take into account the circumstances in light of which she acted, regardless of whether those circumstances actually obtained.<sup>94</sup> Therefore, if Dancy's argument is to be accepted, then it seems that not only are reasons distinct from desires, but they are also motivating without needing to appeal to internal states. This must be the case because the agent herself invokes the object of her beliefs (i.e., states of affairs) as reasons for her action without consideration of her internal belief-states themselves; those states of affairs must have been motivating for her.

## 2.5 The Current Picture

In chapter 1, I argued that substance causation best accounts for our causal observations, thereby demonstrating an ontology of substances. In this chapter, I contend that the human substance is not causally determined by her desires, leaving her free and rational

---

<sup>93</sup> Ibid., 109-110.

<sup>94</sup> Ibid., 110-111.

to originate causal chains. I am partly on my way to demonstrating incompatibilist human free will, but there is still the question of whether the human substance is determined by her ontological makeup, that is, whether her physicality has any causal bearing on her mental choices. Before I can deal with this mind-body problem, however, I need to demonstrate that phenomenal properties exist. Otherwise, we cannot say that there are mental properties that would need to be reconciled with physical ones, if there is indeed this distinction between the mental and physical. Thus, for chapter 3, I turn to the question of qualia and seek to demonstrate the reasonableness of holding its existence.

## 3. The Reasonableness of Qualia

### 3.1 Introduction

This chapter will address the well-trodden question of whether or not qualia exist. This is an important question to cover since in the next chapter, on the mind-body problem, the existence of qualia will be assumed. This chapter, however, will not demonstrate definitively that qualia exist or do not exist, since I do not believe that the existence of qualia can be definitively demonstrated or refuted (a claim about that I hope to support through the course of this chapter). The goal of this chapter therefore is to demonstrate the reasonableness of holding qualia as existent (i.e., non-eliminativable) against those who claim the unreasonableness of holding such a position, in order for the assumption of qualia in the next chapter to be a reasonable one.

I will be demonstrating a cumulative case for the reasonableness of qualia via two lines of argument. First, starting in section 3.3, I discuss epistemic issues surrounding the question of qualia—namely, the treatment of introspection as evidence—in order to clarify the inherent ineffectiveness of arguments against (and, reciprocally, for) qualia. I hope to accomplish two things in this section: 1) demonstrate that eliminativist discussions on qualia rely on begging the question of qualia's non-existence prior to argumentation, as apparent in eliminativists treatment of introspective evidence (reciprocally, pro-qualia theorists are also guilty of begging the question in favor of qualia's existence in a similar manner, as will be discussed); and 2) demonstrate that eliminativists' rejection of introspective evidence does not hold up under the framework within which eliminativists attempt their rejection. The goal of this section is not to positively demonstrate the reality of qualia, but to defend the plausibility of qualia against those making positive arguments against it. At the end of this section, I conclude that, in light of eliminativist arguments against qualia, the question of qualia is left unanswered.

Second, starting in section 3.4, I attempt to provide a positive argument in favor of holding qualia realism. For clarification, throughout this section I will be assuming, without argument, a Kripkean theory of reference. This may be troubling for some philosophers, but rather than arguing for a theory of reference, I will, for expediency's sake, be assuming one. I will also have to assume the authority of introspective evidence. Although this does beg the question of qualia realism, the question will need to be begged in order to get the argument off the ground, and section 3.3 should provide justification for this. Last, I will assume a naturalistic ontology—an ontology thought to be incompatible with qualia realism—and subsequently provide a plausible account of qualia within it. However, as will be apparent in the next chapter, I am not committed to a naturalistic ontology. Nevertheless, if it can be shown that the existence of qualia is compatible with naturalism, then it is shown to be compatible with any ontology.

### 3.2 Defining Qualia

I define qualia as the *qualitative* properties and features of subjective experience. By *qualitative*, I mean those features which, if qualia realism is true, cannot be reduced to, or eliminated by, a qualitative or quantitative description of non-experiential aspects of reality. In other words, qualia are features of consciousness that are not constituted by what we would think of as material or physical mechanistic properties.<sup>1</sup> Hence, we would say that my experience of the color red is indeed my experience of a phenomenal property. Moreover, it is distinguishable from and not eliminated by any description of the mechanical and causal processes behind the generation of my color experience. When I see red, there is something about my experience that is above and beyond the description of causal and mechanistic processes: my phenomenal experience of red.

---

<sup>1</sup> In the classic Cartesian sense of the material or physical as distinct from the mental. There are, of course, contemporary philosophers who take the material and mental to be one and the same. See chapter 4, sections 4.3, 4.4, and 4.5, where I discuss theories that the notion of a distinct physical realm, namely *panpsychism* and *idealism*.

Color is commonly used as the perennial example of qualia (indeed, color qualia is largely used as the focus of this chapter), but qualia includes intrinsic (to the subject's first-person perspective) qualitative content of any kind, such as the taste of garlic, or the sound of a glass breaking, the feel of soft suede, the smell of sulfur, the feel of ambient coldness outside during winter, and the experience of distressing emotional states during traumatic events, amongst other things.<sup>2</sup> Certainly, there are mechanistic and causal processes involved in these experiences; this is not being denied here. What is being denied, however, is that the experience of qualia is wholly constituted by, and thereby eliminativable to, mechanistic processes and properties extrinsic to the subject's first-personal conscious state.

Inherent in this definition of qualia is the *subjective* nature of *experience*, and the *experiential* nature of *consciousness* (indeed, these are what constitute the first-personal conscious state in which qualia operate). These notions are so important to my definition of qualia that, as George Graham and Terence Horgan argue, if there are no such things as qualia, then there are no conscious mental states as we know them, or seem to know them.<sup>3</sup> Consciousness just seems to be constituted by qualia; without qualia, consciousness—or at least, how human beings experience consciousness—could not exist since there would be nothing to constitute it.

### 3.3 The Ineffectiveness of the Qualia Debate

Eliminativism of the phenomenological nature of consciousness in general, and of qualia in particular, attempts to reduce what is seemingly phenomenally experienced by a subject first-personally to what can be mechanically, materially (as matter is thought of by developed theories of the natural sciences), and third-personally explained. A straightforward

---

<sup>2</sup> See also David J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (Oxford: Oxford University Press, 1996), 4.

<sup>3</sup> George Graham and Terence Horgan, "Qualia Realism, Its Phenomenal Contents and Discontents," in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 89.

way of doing this is to just to locate qualia in extended physical space. Such are the theories of Harold Brown, who claims that qualia are displays internal to the brain,<sup>4</sup> and Edward Averill, who argues that phenomenal properties are really just non-dispositional properties of certain cellular activity.<sup>5</sup> Even though these philosophers do not think that qualia can be functionally accounted for (i.e., they affirm that qualia real), they still attempt to account for them as some type of spatially located physical property.<sup>6</sup> However, asserting that qualia is located in the brain or identifying phenomenal properties with certain aspects of the nervous system merely pushes the problem of qualia back an explanatory step, while simultaneously failing to provide any explanation as to why a *subject* would have a seemingly *qualitative experience* in light of purely physical processes.

David Chalmers calls this subjective experience the “key explanandum” in the study of the mind: what needs to be accounted for by any theory of consciousness is the intrinsic nature (or at least seemingly intrinsic nature) of the subjective experience of qualitative sensations—such as when one experiences seeing the color red—and not just what structural mechanisms occur along with experience.<sup>7</sup> This question is difficult for theorists to provide an answer for, however—indeed it is what Chalmers has labeled the “hard problem of consciousness.”<sup>8</sup> One reason for the difficulty, as Paul Churchland points out, is that we have

---

<sup>4</sup> Harold I. Brown, “The Case for Indirect Realism,” in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 47-49, 53-54, 56.

<sup>5</sup> Edward Wilson Averill, “The Phenomenological Character of Color Perception,” *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 157 (2012): 32-33.

<sup>6</sup> Brown, “The Case for Indirect Realism,” 45-46; Averill, “The Phenomenological Character of Color Perception,” 30.

<sup>7</sup> David J. Chalmers, “Facing Up to the Problem of Consciousness,” *Journal of Consciousness Studies* 2 (1995): 206, 213.

<sup>8</sup> *Ibid.*, 201.

*prima facie* reason to believe in qualia realism based on the first-person introspection of our own experiences.<sup>9</sup>

Therefore, the problem with determining the reality or unreality of qualia lies with the problem concerning what to do with evidence gained from introspection—in fact, determining the ontological status of qualia seems solely dependent on how this evidence is handled. This is because, if qualia are real, the only way qualia are observed is first-personally. In other words, evidence for the existence of qualia can *only* be gained via first-personal observation—i.e., introspection—if it can be gained at all. Evidence affirming qualia’s existence seems to rely solely the knowledge that can be gained via introspection and only introspection. Likewise, evidence against eliminativism arrives solely via the avenue of introspection.

Is the evidence gained via introspection even epistemically valuable as it is presented, so to speak, or can we dismiss it? Because introspective evidence is the only evidence for qualia, answering this question one way or another begs the question of whatever position the answer favors. That is, if we take first-personal evidence as authoritative, then our deliberations will favor the conclusion of qualia realism. On the other hand, if we dismiss the authority of first-personal evidence, then our ruminations will be skewed towards eliminativism. This is because determining whether to dismiss introspective evidence requires evaluating the contents of introspection. However, evaluating the contents of introspection is the primary goal of inquiry to begin with. In this section, I seek not to evaluate the contents of introspection, but to analyze how various eliminativist theorists have attempted to dismiss the contents of introspection as evidence. I aim to show that attempts to dismiss such evidence have failed to be justified.

---

<sup>9</sup> Paul M. Churchland, *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science* (Cambridge, MA: MIT Press, 1989), 25.

### 3.3.1 Tactics for Dismissing Introspective Evidence

#### 3.3.1.1 *Collapsing First-Person to the Third-Person*

One way of dealing with first-personal evidence is to reduce it to third-personal evidence, thereby eliminating the problems that theorists face when accounting for the seemingly phenomenally intrinsic nature of first-personal introspective evidence altogether. This is the tactic of Paul Churchland, who argues that phenomenal experience just is the subject's experience of her own brain states. In other words, the subject gains knowledge of her own brain states when she is introspectively aware of her experiences.<sup>10</sup> For Churchland, the content of subjective introspection just is what is predicated by neuroscience in descriptions of physical brain states.<sup>11</sup> Likewise, the subject's supposedly phenomenal experience of the color red just is her coming into contact with a surface of an object that reflects a certain wavelength of light.<sup>12</sup> Thus, any phenomenal introspective experience is accounted for via objective, physical processes and features of the subject's brain, along with her environment. If we accept this definition of introspective experience, then what seems to be *sui generis* first-personal accounts of internal states *just are* third-personal accounts of physical processes.

We can clarify Churchland's first-person/third-person collapse by looking at his claims. To begin, Churchland argues that it is possible great amounts of knowledge concerning physical states can be gained through introspection:

I suggest, then, that those of us who prize the flux and content of our subjective phenomenological experience need not view the advance of materialistic neuroscience with fear and foreboding. Quite the contrary. The genuine arrival of a materialist kinematics and dynamics for psychological states and cognitive processes will constitute not a gloom in which our inner life is suppressed or eclipsed, but rather a dawning, in which its marvelous

---

<sup>10</sup> Ibid., 60, 64-66.

<sup>11</sup> Ibid., 54-55.

<sup>12</sup> Ibid., 56.

intricacies are finally *revealed*—most notably, if we apply ourselves, in direct self-conscious introspection.<sup>13</sup>

If introspection gains knowledge of physical processes like this, then it is capable of gaining empirical evidence, and therefore is a third-person vantage point. But whether Churchland is indeed collapsing the first-person/third-person distinction is still unclear. However, later Churchland is explicit concerning the nature of what we consider “first-person” accounts. He states:

The intricacies of brain function may be subjectively opaque to us now, but they need not remain that way forever. Neuroscience may appear to be defective in providing a purely “third-person account” of mind, but only familiarity of idiom and spontaneity of conceptual response are required to make it a “first-personal account” as well. *What makes an account a “first-person account” is not the content of that account, but the fact that one has learned to use it as the vehicle of spontaneous conceptualization in introspection and self-description* [Churchland’s italics].<sup>14</sup>

The idea that the first-person vantage point consists in “only familiarity of idiom and spontaneity of conceptual response,” and the fact that Churchland implies that these are unnecessary for description, is indicative of how little the first-person perspective is valued according to Churchland’s theory. But can the first-person vantage point be done away with and replaced with the third-person vantage point? According to Churchland, it can and should. Since first-personal “spontaneous conceptualization” is so misleading as descriptions of physical processes, Churchland advocates that we change the way human beings discuss their psychological and mental states:

Old ways die hard, and in the absence of some practical necessity, they may not die at all. Even so, it is not inconceivable that some segment of the population, or all of it, should become intimately familiar with the vocabulary required to characterize our kinematical states, learn the laws governing their interactions and behavioral projections, acquire a facility in their first-person ascription, and displace the use of [folk psychology, i.e., the way of communicating that presumes the reality of phenomenal experience]

---

<sup>13</sup> Ibid., 66.

<sup>14</sup> Ibid., 74-75.

altogether, even in the marketplace. The demise of [folk psychology's] ontology would then be complete.<sup>15</sup>

A change in language use would equate to the end of first-person ascriptions, thereby making the collapse of the first-person into the third-person complete.

This collapse, however, is only justified if it can be said that the content of first-personal introspection is indeed physical processes. But is it? Churchland does specifically address this problem via a discussion of qualia, since qualia are quintessentially considered the content of first-personal introspection. However, due to his eliminativist project, he defines qualia as such:

Qualia...are not an ineffable mystery, any more than colors or temperatures are. They are physical features of our psychological states, and we may expect qualia of some sort or other in any physical system that is sufficiently complex to be functionally isomorphic with our own psychology. The qualia of...a robot's states are not "absent." They are merely *unrecognized* by us under their physical/electronic descriptions, or as discriminated by the modalities of outer sense rather than inner sense.<sup>16</sup>

Churchland goes on to claim that the existence of qualia need not be denied, as long as qualia are thought of as fulfilling his physical definition;<sup>17</sup> thus, according to Churchland's theory, qualia are extrinsic, scientifically quantifiable properties and third-personally observable.

The question remains, however, whether Churchland's reductive redefinition of qualia is warranted. It seems that it is not unless we beg the question against qualia realism in the classic sense. Churchland does address this issue while considering an objection concerning qualitative subjective state *kinds*, specifically pain states that are reported and observed correlating with a variety of physical states. He admits that it seems unlikely that animals' various functional pain states share a common physical

---

<sup>15</sup> Ibid., 18-19.

<sup>16</sup> Ibid., 31.

<sup>17</sup> Ibid.

nature. However, he maintains that, given evolutionary theory, dualism is implausible. Therefore, given the empirical evidence, we should presume against natural kinds of psychological states. To give into our intuitions otherwise begs the question against functionalism.<sup>18</sup>

What seems unadmitted here, however, is that qualia realism as classically thought of is not just exclusive to dualism, but is also compatible with non-reductive physicalism, panpsychism, and monism. Even if evolutionary theories are incompatible with dualism (which has yet to be firmly established), it has yet to be shown that they are exclusively compatible with reductivist physicalism. But more to the point, to presume *against* psychological state kinds may be presuming so much that it is begging the question of eliminativism, especially without providing an evolutionary theory that firmly disallows qualia. This is because empirical evidence can only be used against introspective evidence if first-personal introspective evidence was of an empirical (i.e., scientifically quantifiable) nature. But classic qualia realism does not hold this. To argue that empirical evidence presumes against classic qualia realism is just to state a fact of presumption, not to provide an argument against classic qualia realism based on this evidence. In agreement with Churchland, I admit that to define qualia classically is begging the question of qualia realism, but to redefine qualia on the basis of empirical presumption is begging the question the other way around.

Furthermore, Churchland's reduction of first-personal vantage points to third-personal vantage points does not do justice to the question of why mental states seem to be private. I can experience the seemingly phenomenal content, either reducible or nonreducible, of my own brain states, but you cannot. You can only experience the non-phenomenal, mechanistic

---

<sup>18</sup> Ibid., 25-26.

brain states themselves. But if we are both experiencing this same event third-personally, we should have similar experiences of the event. But we do not. I argue that this is because my experience of the color red for you and for me are experiences of *two different things*.

Therefore, my seemingly phenomenal experience of the color red and my brain states cannot be identical to each other.<sup>19</sup> If they were, then you should be able to have the seemingly phenomenally experience my brain states just as I can. But as it is, I experience what seems to be qualia of the color red, while you observe certain of my brain cells firing. In fact, you having a seemingly phenomenal experience of the color red as an awareness of my brain states seems absurd. Thus, the first-personal vantage point seems as if it cannot be identical to the third-personal vantage point.

In support of Churchland's view, one may counter that, in the above scenario, my vantage point relative to my brain determines the nature of my third-personal experience of my own brain states. Someone else who is having an experience of my brain states is merely experiencing them from a different vantage point. Churchland seems to acknowledge this when he makes reference to "the modalities of outer sense" and "inner sense" in the quote above.<sup>20</sup> I would agree with this, that these different vantage points equate to different experiences, but add that the vantage point one has in relation to their own brain is the *subjective* position, and that this position's experience, as such, is *first-personal*. This is what gives the position the markedly different character that it has from the third-personal position. And as such, we should consider the first-personal vantage point differently than the third-personal, and not presume that the knowledge that can be gained first-personally is identical to the knowledge that can be gained third-personally.

---

<sup>19</sup> See also Mark Crooks, "The Churchlands' War on Qualia," in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 207.

<sup>20</sup> Churchland, *A Neurocomputational Perspective*, 31.

### 3.3.1.2 *Negating the First-Person*

Another other way of dealing with first-personal introspective evidence is to devalue it, thereby negating the need to consider it. Such is Daniel Dennett’s approach. Unlike Churchland, Dennett doesn’t assume that introspection results in valuable third-personal knowledge of brain states. Instead, introspection for Dennett seems an exercise that bears little epistemic fruit:

[T]here is yet another possibility [concerning the nature of phenomenological theorizing], which I think is closer to the truth: what we are fooling ourselves about is the idea that the activity of “introspection” is *ever* a matter of just “looking and seeing.” I suspect that when we claim to be just using our powers of inner *observation*, we are always actually engaging in a sort of impromptu *theorizing*—and we are remarkably gullible theorists, precisely because there is so *little* to “observe” and so much to pontificate about without fear of contradiction.<sup>21</sup>

His argument hinges on our tendency to be wrong about the content of our perceptions, a fact which he hopes succeeds in undermining the value of taking any first-person account of experience gained via introspection as evidence for phenomenal realism.<sup>22</sup> To support this, Dennett points to the fact that, although we are confident about it, we are often wrong about the content of our peripheral vision. Moreover, Dennett argues, the overconfidence we have concerning this content arises from a misconception about what our visual field is, i.e., we think it is “an inner picture composed of colored shapes.”<sup>23</sup> Instead, we should lack confidence in our subjective states.

But what about the notion of privileged access? Dennett does not outright deny that we have some measure of privileged access. Nevertheless, he argues that the knowledge of one’s own subjective states lacks the epistemic authority we tend to think it possesses.<sup>24</sup>

---

<sup>21</sup> Daniel C. Dennett, *Consciousness Explained* (London: Penguin Books Ltd, 1991), 68-69.

<sup>22</sup> *Ibid.*, 68-70.

<sup>23</sup> *Ibid.*, 68.

<sup>24</sup> *Ibid.*, 68-69.

Since we are often wrong about the content of our subjective states, we should have little confidence in introspection revealing the true nature of those states. Thus, we cannot claim any knowledge concerning the true nature of those states based on introspection alone.<sup>25</sup>

Therefore, in lieu of introspection, Dennett affirms the authority of the third-person perspective in giving an account of subjective experience. Thus, Dennett argues, phenomenological experience should be analyzed utilizing what he calls the “heterophenomenological method.”<sup>26</sup> This method, which heavily relies on the third-person standpoint, is intended to be the scientific and “neutral” way of exploring the subjective experience of consciousness, and begins by making no assumptions concerning consciousness—including the assumption that human beings have consciousness full of content.<sup>27</sup> Part of this method requires that we ascribe to the subject an intentional stance, that is, when receiving a first-personal account of a subject’s experience, all that we should assume about the subject is that she has intention behind her speech.<sup>28</sup> With these parameters in mind, the question of human consciousness can be explored.

Under these parameters, Dennett then addresses evidence concerning the seeming phenomenal quality of consciousness—vision qualia specifically—as it seems to exist in the “Cartesian Theater” of the mind (i.e., the place where phenomenal experience is supposedly played out for a subject to witness, as Dennett himself characterizes mental and subjective realism).<sup>29</sup> Dennett points to an instance of “prosthetic vision,” an experiment where a blind subject was given “sight” via a pixel mechanism which tactilely reproduced what would be in

---

<sup>25</sup> Ibid., 70-71.

<sup>26</sup> Ibid., 72.

<sup>27</sup> Ibid., 72-73.

<sup>28</sup> Ibid., 76.

<sup>29</sup> Ibid., 39.

the subject's line of vision on his back or chest. With this mechanism, subjects could read signs and recognize objects and faces, much like a sighted person.<sup>30</sup> Concerning the blind person's ability to gain information of the environment around him via this form a "sight," Dennett questions whether the phenomenal qualities of the subject's experience was that of vision or merely touch sensation.<sup>31</sup> Dennett concludes that, since the blind subject had much of the same experience as a sighted subject (including exhibiting the same behaviors a sighted person would exhibit if they were experiencing vision qualia), information gained of the outside world is all there is to experience. If tactile "vision" could transmit as much information of the exterior world as regular vision, then the notion of vision qualia can be completely dismissed.<sup>32</sup> This conclusion concerning information negating the need for qualia is also evident in his treatment of knowledge gained by perception and its relation to any supposedly "phenomenal" content of consciousness. Take for example Dennett's thought experiment of Marilyn Monroe wallpaper.<sup>33</sup> Dennett argues that it makes sense if a visually functioning subject were to develop the belief that a wall covered in Marilyn Monroe wallpaper is indeed covered in Marilyn Monroe wallpaper, for the wall *is* covered in Marilyn Monroe wallpaper. This does not mean, however, that the image of Marilyn Monroe wallpaper was recreated in the subject's mind.<sup>34</sup> By this argument, Dennett is assuming that, since there is no need to postulate any phenomenal content of consciousness in order to explain the knowledge concerning the wallpaper gained via experience, there is no need to

---

<sup>30</sup> Ibid., 339-340.

<sup>31</sup> Ibid., 340.

<sup>32</sup> Ibid., 338-343.

<sup>33</sup> Ibid., 354.

<sup>34</sup> Ibid., 364.

postulate phenomenal content, along with qualia, at all. Our inner “seemings” in fact do not exist.

With his notions of heterophenomenology and the intentional stance, and his downplaying of introspection, Dennett concludes that qualia do not exist. This treatment of introspection, of course, begs the question at issue. That is, if Dennett has no independent reason to dismiss evidence from introspection, then he’s unjustified in doing so, since to dismiss this evidence requires evaluating the contents of introspection, and evaluating this content is what the whole argument is about. The question then becomes, has Dennett provided a good independent reason for treating introspection in the way that he does? As discussed above, Dennett seems to offer two: 1) subjective accounts of experience are prone to error and 2) there is no need to postulate qualia when information flow/knowledge is accounted for without it. However, even though Dennett makes a good case for propositional error, it does not follow from Dennett’s argument that a subject is not experiencing *something* and, further, that this something is not *phenomenal*.<sup>35</sup> The subject may still be having a phenomenal experience, however incorrect they are about that experience. It doesn’t follow that just because we cannot be confident about the *contents* of our experience that we cannot be confident that we are undergoing *an experience*.

Likewise, in the case of information flow/knowledge, it does not follow that knowledge gained of the outside world negates the existence of qualia. For instance, in the case of the “prosthetic vision,” it does not follow that change in qualitative kind—from visual to tactile—results in the negation of qualia altogether if “vision” is proven to be informative. In fact, this seems to demonstrate that information is lost if

---

<sup>35</sup> See also Galen Strawson, *Mental Reality*, 2<sup>nd</sup> ed. (Cambridge, MA: MIT Press, 2009), 52.

we ignore qualia, considering the experience of tactile vision and regular vision must be internally very different for the subject experiencing it. Moreover, even with perfect knowledge gained, Dennett cannot escape the problem of phenomenology. According to Dennett himself, it *seems* as if phenomenological experience exists, even though it doesn't.<sup>36</sup> The *seemingly phenomenological* nature of experience, however, is itself phenomenological and first-personal. Thus, the qualitative nature of experience itself cannot be reduced with the reduction of the mere content of experience.<sup>37</sup> It must also be accounted for. If this phenomenological seeming surrounding experience itself exists, we are still burdened by unreduced first-personal introspective evidence.

### ***3.3.1.3 Redefining the Content of First-Personal Experience***

It seems as if a brand of eliminativism that attempts to collapse first-personal evidence into third-personal, or attempts to ignore first-personal evidence altogether, inadequately addresses the problem of first-personal introspective evidence. However, there is another approach: redefine the content of first-personal experience to begin with, and base the interpretation of first-personal observation and evidence on said definition. Such is the aim of representationalism, a view of consciousness which attempts to redefine phenomenal experience in terms of intentionality.<sup>38</sup>

Broadly speaking, there are two types of representationalist approaches: one weak and one strong. The weak holds that mental states are universally intentional, although some

---

<sup>36</sup> Dennett, *Consciousness Explained*, 366.

<sup>37</sup> See also Strawson, *Mental Reality*, 52, for a similar assessment of Dennett's claim.

<sup>38</sup> John O'Dea, "Transparency and the Unity of Experience," in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 299; Also note, in this chapter I focus the discussion of representationalism on the ontological status of mental content. In chapter 7, I discuss representationalism again, but in the context of a discussion on the possibility of non-conceptual perceptual knowledge. Hence, the topic of perceptual knowledge is addressed in this chapter since it has bearing on the ontological status of the content of mental states, but a more thorough theory of perceptual knowledge will be presented in Chapter 7.

content of those states can be non-intentional. On the other hand, the strong approach strictly affirms the universal intentionality of all mental content in all mental states.<sup>39</sup> How is it that the representationalist can affirm this intentionality in light of phenomenal experience? As Howard Robinson puts it, representationalism rests on the notion that experience is *transparent*, that is, “there is nothing introspectable in experience other than the external.”<sup>40</sup> This characterization of experience makes representationalism at heart an eliminativist project, since the goal is to eliminate the need to postulate the existence of phenomenal intrinsic experience.<sup>41</sup>

One such theory is that of Fred Dretske. Dretske argues that the quality of a subject’s experience should be distinguished by its intentionality, that is, by the relation between subject and the intentional objects of that experience, in the same way a subject’s belief is distinguished by its intentional object. Under this assumption, Dretske rejects the idea that experiences are constituted by qualia in the mind of the subject.<sup>42</sup> This rejection stems from an inability to situate the qualities that individuate one experience from another within the subject. Dretske reasons that these qualities therefore must be relational, and indeed intentional.<sup>43</sup> Given this, Dretske rejects local supervenience (e.g., supervenience on brain states) and claims that what we would consider qualitative properties—such as *red* or *hot* and

---

<sup>39</sup> Barry Maund, “A Defense of Qualia in the Strong Sense,” in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 272.

<sup>40</sup> Howard Robinson, “Why Frank Should Not Have Jilted Mary,” in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 227.

<sup>41</sup> Martine Nida-Rümelin, “Phenomenal Character and the Transparency of Experience,” in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 311. See also Amy Kind, “How to Believe in Qualia,” in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 285.

<sup>42</sup> Fred Dretske, “Phenomenal Externalism or If Meanings Ain’t in the Head, Where Are Qualia?” *Philosophical Issues* 7 (1996): 144-145.

<sup>43</sup> *Ibid.*

the like—are outside the subject’s mental states and in a relation with the objects that mental states are experiences of.<sup>44</sup>

In support of his theory, Dretske argues that people experience the objects of their experience through their beliefs about their experience. In other words, what I believe about the objects of my experience determines the quality of the experience that I have of those objects. To support his argument, he distinguishes between two different meanings of the term “look” (as in “look,” “seem,” and “appear”): the phenomenal sense, which picks out the phenomenal quality of how objects present themselves to a subject undergoing an experience, and the doxastic sense, which picks out the beliefs held by a subject that bears upon the experience in question.<sup>45</sup> The point Dretske makes is that a way an object appears to a subject in the phenomenal sense is determined by the way it appears in the doxastic sense, since the subject’s beliefs about her experience determines the quality of experience.<sup>46</sup> To demonstrate the truth of this thesis, Dretske utilizes Putnam’s thought experiment of Fred and Twin Earth Fred (i.e., “Twin Fred”) both encountering *water* (i.e., as opposed to Twin Earth “twater”) and the content of their individual experiences. For clarification, Dretske utilizes “LOOKS” in all capitals to indicate “look” in the phenomenal sense, and “*k*” is a puddle of water. Dretske argues:

Fred will think—correctly as it turns out—that his experience of the puddle (the way the puddle LOOKS to him) is exactly like his experience of water (is exactly the way water normally LOOKS to him). So Fred will think—once again correctly—that *k* LOOKS like water to him. Twin Fred, on the other hand, will think that *k* LOOKS like twater to him. He will think—correctly—that *k* LOOKS the way twater always looked to him in normal conditions. The twins’ phenomenal experience of the water will prompt in them, not only different beliefs about the water, but different beliefs about their phenomenal

---

<sup>44</sup> Ibid.

<sup>45</sup> Ibid., 147-148.

<sup>46</sup> Ibid., 149-150.

experience of it. How their experience seems to them will therefore be different.<sup>47</sup>

Dretske likens this difference in phenomenal experience based on doxastic considerations to the experience of a small child with a petunia and the way Dretske's own experience of a petunia would be. Because the child has no knowledge of a petunia while Dretske has a developed concept of a petunia, Dretske's experience of a petunia would seem to Dretske phenomenally different than a child's experience of a petunia would seem to her.<sup>48</sup> Because of this difference, doxastic considerations must bear on phenomenal experience.

Susanna Schellenberg, in espousing her own representationalist theory, makes a similar kind of argument:

If I speak Urdu, then a sentence uttered in Urdu sounds different to me than if I do not speak Urdu. Similarly, if I possess the concept of a skyscraper, then a tall building arguably looks different than if I do not possess the concept. If experience is argued to have [representational] content, then these differences can easily be accounted for.<sup>49</sup>

In addition to the above, Schellenberg offers the following positive argument in favor of conscious experience consisting in representational content:

Premise 1: If a subject is perceptually related to the world (and not suffering from blindsight etc.), then she is aware of the world.  
Premise 2: If a subject is aware of the world, then the world seems a certain way to her.  
Premise 3: If the world seems a certain way to her, then she has an experience with content *C*, where *C* corresponds to the way the world seems to her.  
Conclusion: If a subject is perceptually related to the world (and not suffering from blindsight etc.), then she has an experience with content *C*, where *C* corresponds to the way the world seems to her.<sup>50</sup>

---

<sup>47</sup> Ibid., 150.

<sup>48</sup> Ibid., 148-149.

<sup>49</sup> Susanna Schellenberg, "Perceptual Content Defended," *Nous* 45 (2011): 719.

<sup>50</sup> Ibid.

In support of the third premise, Schellenberg proposes what she calls the “seems-content link,” that is, the natural connection between a subject’s perception and its contents.<sup>51</sup> This link allows that any given object of perception can be perceived in a variety of ways, since more than one set of contents can correspond to any given object of perception. As Schellenberg argues:

The way the world noncomparatively seems to the perceiver may change from moment to moment even as her gaze remains steady. Say she is looking at a pig. She can direct her attention at its shape, its color, the texture of its skin, or any combination of these features. As her attention shifts, her phenomenology will change. One or more propositions can be associated with every one of these phenomenal states and thus with every one of these ways that the world may noncomparatively seem to her.<sup>52</sup>

In this way, utilizing Dretske’s example, the adult Dretske can have a phenomenally different experience of a flower than a small child. Such is also with the speaker and non-speaker listening to Urdu being spoken. The seems-content link explains why these differences in experiences exist.

Schellenberg also points to our ability to gain knowledge through perception as supportive of her conclusion. She presents the thought experiment of a subject perceiving a coffee cup; the object of the cup is what grounds the knowledge of the cup, not the qualitative features of perception. In order to clarify this point, Schellenberg offers up a contrasting case: a subject perceives a coffee cup and then, after closing her eyes a moment, perceives a qualitatively identical but numerically different coffee cup. The subject’s belief that the cup is the same, based on the quality of perception, is false. The truth or falsity of her belief is grounded in the cup—the object of her perception—not in the perception itself.<sup>53</sup>

---

<sup>51</sup> Ibid., 723.

<sup>52</sup> Ibid., 724.

<sup>53</sup> Ibid., 735.

*Ergo* qualia play no role in the generation and grounding of perceptual knowledge, and therefore qualia do not exist.

Does representationalism adequately dismiss the burden that first-personal introspective evidence puts on explicating a theory of consciousness? If we accept Schellenberg's positive argument in favor of representationalism, then it seems as if it does. But, as seen with Dennett's view, Schellenberg's argument also begs the question of eliminativism. This can be clearly seen when analyzing the wording of her second premise. If we replace the representationalism favoring "then the world seems a certain way to her" with the qualia realism favoring "then the world *appears* a certain way to her," then we do not walk away with a valid argument in favor of representationalism. From this, we can see that the word "seems," as used in the second premise by Schellenberg, is in the doxastic sense.<sup>54</sup> Thus, through her choice of words, Schellenberg begs the question of eliminativism. This might seem like a fair move on Schellenberg's part since to utilize "seems" in the phenomenal sense is to beg the question of qualia realism. We must be given more reason to believe that begging the question in favor of eliminativism is more reasonable than the alternative, which Dretske and Schellenberg attempt to do.

Dretske's and Schellenberg's views, outside the positive argument Schellenberg provides, rests on the notion that phenomenology is dependent on doxastic content. They base their argument on the evidence of our own experiences. However, a reasonable case can be made that phenomenology is independent of doxastic content, also based on our experiences. Take Schellenberg's example of Urdu: we can make the case that the speaker and non-speaker of Urdu, the phenomenal experience of hearing someone speak Urdu *sounds* exactly the same. The only difference between the speaker and non-speaker is that the speaker is able to link socially established shared meaning to the auditory forms that he

---

<sup>54</sup> Dretske, "Phenomenal Externalism," 148.

perceiving, while the non-speaker is not. The same is true of Dretske and the child. They are both having the same phenomenal experience, but Dretske is able to connect this current qualitative experience with certain kinds—that is, the *petunia* kind—that he’s had in the past. The phenomenal content has not changed at all, regardless of the doxastic content. It always presents itself the same; the difference is in the subject’s reaction to it.

Moreover, the idea that phenomenology is independent of doxastic content has the potential to make better sense of human learning than the alternative. If we take a dependency relation between phenomenology and doxastic content seriously, then learning via experience seems to be an accidental exercise, since learning could only be cashed out in terms of doxastic content changing on its own and our subsequent experiencing the now new and different content. But what would cause this change of content in the first place? One may argue that a new stimulus has been introduced. But what about when children learn about different features of objects that have not physically changed? The phenomenology-doxastic content dependency does not seem to make sense in these cases. If phenomenology is independent, however, then we can cash out learning as the process of changing how we *attend* to our phenomenology, and our being able to make connections between past and present phenomenological presentations. In reference to Dretske’s example, Dretske knows the phenomenal kind *petunia* because he has experienced the phenomenal kind *petunia* in the past. Dretske has attended to different features of this phenomenal kind *petunia*, and therefore he is able to recognize these features when they present themselves again. The small child has not come across this phenomenal kind in the past, however, and thus, unlike Dretske, lacks the skill to recognize the characteristic features of *petunia*; furthermore, the child is yet able to apply knowledge of past experiences of *petunia* kind to current ones.

Moreover, in regards to the thought experiment of the coffee cup, Schellenberg is wrong to suggest that qualia realism, unlike representationalism, fails to provide a reliable

account of perceptual knowledge. Remember in the thought experiment, the subject perceives a coffee cup and then, after having closed her eyes a short while, perceives a qualitatively identical but numerically different coffee cup. The subject then forms the false belief, based on the quality of her experience, that the coffee cup is the same. Therefore, Schellenberg concludes, we should not ascribe to qualia realism if we want to ascribe to the notion that knowledge can be gained perceptually: dependence on qualia for perceptual knowledge would lead to the formation of false beliefs.

I question, however, how a true belief could be arrived at under the circumstances of Schellenberg's thought experiment, even given representationalism. Schellenberg is right that the truth or falsity of a belief about the coffee cup is grounded in the cup; a qualia realist does not have to deny this. However, under the circumstances of being in the presence of a qualitatively same but numerically different coffee cup, a false belief would be justifiably formed, even in a representationalist ontology. The subject would have no reason to believe it was a different cup, and therefore would lack justification for believing so. In any case, Schellenberg argues that a qualia realist would have to give an account of how a subject would know a coffee cup is the same as a coffee cup she saw at a different time, given that all she has to go by is the phenomenal perception.<sup>55</sup> However, there is nothing according to my version of qualia realism that bars a subject from bringing in background knowledge into an instance of forming perceptual knowledge. In other words, it is compatible with the theory being developed here that background knowledge has a bearing on the contents of perceptual knowledge. What is being denied however is that background knowledge has a bearing on the contents of perception itself.<sup>56</sup>

---

<sup>55</sup> Schellenberg, "Perceptual Content Defended," 736.

<sup>56</sup> These thoughts on perception, background knowledge, and perceptual knowledge formation are discussed much further in Chapter 7.

### 3.3.2 Assessment

If we had independent reason to maintain the reduction, elimination, or redefinition of first-personal introspective evidence, then we would have reason to dismiss first-personal evidence when it comes to the question of qualia. However, as I have attempted to demonstrate above, we do not have independent reason without first begging the question against qualia, since qualia, if they exist, are what would constitute first-personal introspective evidence.

Churchland's collapse of the first-personal vantage point to third-personal vantage point seems too hasty, and still leaves a characterizing feature—namely privacy—of the first-personal vantage point left unaccounted for. Dennett, in his attempt to eliminate the need to take first-personal evidence into consideration, did not adequately give an account of why we should remove it from the equation. Although he independently demonstrated that subjects tend to be in error concerning the content of their experience, this does not disprove the reality of subjective experience altogether. Even if the subject is wrong about every aspect of his experience, the subject is nevertheless undergoing *an* experience. Moreover, the seemingly phenomenological nature of experience is itself phenomenological, and is not eliminated by reducing the contents of experience.

Representationalism attempts to redefine the content of introspection as the representational content of intentional objects, thereby removing the need to posit the existence of any intrinsic qualitative properties, and making all introspectable properties properties of external objects. This move, if justified, would dismiss first-personal introspective evidence as evidence of qualia realism. However, the only reason we might make this move is because it would give us a better account of our experiences than qualia realism. But in some ways, such as in the explanation of learning as discussed above, it seems to fail in this regard.

Amy Kind, in her assessment of representationalism, points out that the representationalist logically errs in holding that the existence of representational content translates into attending consisting in *nothing but* attending to representational content.<sup>57</sup> However, as Kind states, “there is a difference between saying that introspecting an experience of pain *involves* or even *requires* attending to a particular bodily location and saying that *all that there is* to introspecting an experience of pain is attending to a particular bodily location.”<sup>58</sup> Kind is quick to point out that the representationalist case concerning representation becomes muddled in light of emotions: what external object exactly is the subject attending to when the subject attends to an emotion? There isn’t an easy representationalist answer.<sup>59</sup> We could posit that in the case of emotions the subject is attending to a wide variety of objects at once. But although this seems to be a reasonable avenue of exploration, it does highlight some problems for the representationalist to address.

Nida-Rümelin also makes a similar point as Kind, claiming that the representationalist assumes we cannot be aware of the intrinsic properties of our own experiences on the basis that we can be aware of the properties of external objects. Nida-Rümelin, however, points out one assertion does not follow from the other—it is feasible we can both be aware of the properties of external objects and the intrinsic properties of our own experience,<sup>60</sup> as my answer to Schellenberg’s coffee cup problem indicates. Diana Raffman further suggests that it is unreasonable that we could be aware of any content of experience *without* being aware of intrinsic qualitative properties of experience to begin with.<sup>61</sup> In any case, if we are aware of

---

<sup>57</sup> Amy Kind, “How to Believe in Qualia,” 294.

<sup>58</sup> *Ibid.*, 292.

<sup>59</sup> *Ibid.*, 293.

<sup>60</sup> Nida-Rümelin, “Phenomenal Character and the Transparency of Experience,” 314-315.

<sup>61</sup> Diana Raffman, “From the Looks of Things: The Explanatory Failure of Representationalism,” in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 328.

intrinsic content in any capacity, even if some form of representationalism is true, this would count as first-personal introspective evidence in favor of qualia realism, and we would still be left where we started, trying to evaluate the importance of first-personal introspective evidence.

The point of section 3.3 is not to make a demonstrative case against eliminativism, but to defend first-personal introspective evidence as evidence for qualia realism. However, I recognize that maintaining first-personal introspective evidence is evidence for qualia realism begs the question of qualia realism. Thus, all that I hope to have accomplished in this section is a defense of the reasonableness of taking first-personal introspective evidence as evidence of the existence of qualia. What follows in section 3.4 is more of a positive independent argument for the existence of qualia within a naturalistic framework. However, the argument still relies on evidence gained from the first-personal point of view. My hope in section 3.4 therefore is not to definitively demonstrate the existence of qualia, but to give some plausibility to the belief that qualia exist. And if qualia are plausible in a naturalistic ontology, then they can be considered plausible in any ontology.

### **3.4 Positive Argument for The Existence of Qualia**

Saul Kripke famously argued for the non-identity the mental and physical utilizing the example of the mental state of *pain* and the physical brain state of *C-fibers firing*. Pain cannot be identical to C-fibers firing, Kripke argues, because it is possible for someone to be in what we would call—that is, what we have *rigidly designated*—as *pain* without C-fibers firing. Since this is possible, an identity relation (i.e., a relation of necessity) cannot exist between the experience of pain and C-fibers firing.<sup>62</sup>

---

<sup>62</sup> Saul Kripke, *Naming and Necessity* (Malden, MA: Blackwell Publishing, 1981), 98-99.

The notion of rigid designation is the notion of picking out a referent, as it essentially is, across all possible worlds. Along with this notion comes the notion of necessary properties; an individual person or object has a set of necessary properties without which that person or object fails to be itself and therefore does not exist. These are different than contingent properties, which a person or object can possess or fail to possess without impact on its being identical to itself and in a state of existence.<sup>63</sup>

Kripke argues that picking out a person or object specifically is born out of people's common reference to contingent properties. That is, attempts at *de re* reference *forces us into making contingent predication*. An example Kripke uses is an attempt to distinguish Richard Feynman, the famous 20<sup>th</sup> century physicist, from a random unassuming person named Feynman. In both cases, the rigid designator *Feynman* can be utilized to pick out both Feynmans, since they are both indeed, in a *de re* sense, identical to Feynman (that is, each *Feynman* is identical to itself). However, even though they are both necessarily identical to a Feynman they are not identical to the *same* Feynman. How do we make a distinction then between the two Feynmans? We'd have to go beyond the rigid designator in reference to Feynman. To pick out the famous physicist, we'd need to make reference to his being a *famous physicist*. His being a famous physicist, however, is *not an essential* property of the famous Feynman: it is a *contingent* one, since he could have failed to become a physicist.<sup>64</sup>

With the above in mind, a challenge for us when we are exploring the issue of identity between the mental and the physical is to distinguish necessary from contingent properties. Taking seemingly qualitative properties as our paradigmatic example of a possibly intrinsic property, I will defend the claim that either the way a color appears is a necessary property of color while the physical properties of color are contingent properties,

---

<sup>63</sup> Ibid., 3-21, 46-58.

<sup>64</sup> Ibid., 80-81.

or the way a color appears is a contingent property of color while the physical properties of color are necessary properties. Therefore, the relation between appearance and any physical property is a contingent one, thereby undermining any eliminative identity thesis. This would be against those like Churchland and Dennett who would argue that the experiential components of color are in a necessity relation to physical processes found in the human body, or those like Dretske and Schellenberg who would argue that these experiential components are in a necessity relation with intentional objects.

So let us discuss the necessary and contingent properties of color. Suppose there existed a possible world where the visible light spectrum wavelengths we in the actual world would associate with the color *green* instead appeared to us humans as what we in the actual world would call *red*, and vice versa (i.e., we evolved to see colors inverted).<sup>65</sup> How would we be able to uniquely identify the color *red* across these worlds? In other words, given both the actual world and this possible inverted color world, what entity does the word *red* rigidly designate? What are its properties? In order to answer this question we must make a distinction between possible world *red* and actual world *red* in the same way we make a distinction between famous physicist *Feynman* and some other *Feynman*. That is, we must make reference to the color *red*'s *contingent properties*.

What are *red*'s contingent properties? It depends on what you take *red*'s necessary properties to be, since whatever property of *red* is not necessary is contingent. If you take *appearing actual world red to humans* to be a necessary property of the color *red*, then any appearance of actual world *red* to humans, regardless of physical grounds, will itself be *red*. Thus, in the possible world in question, leafy plants and trees will be *red* since they appear to

---

<sup>65</sup> Let us say that in this possible world, this inversion is a product of evolution; human vision just evolved this way in this world in response to this world's human beings' unique environmental needs. Perhaps leafy plants were scarcer, of more nutritional value, and, as a consequence, were of more survival interest for humans on this world. Therefore, being able to spot them more easily was of utmost importance to our species' success. On the other hand, berries and fruit were unimportant, and therefore humans had no need to spot them as easily.

humans as *red* appears in the actual world. Given this (that is, if we take the *appearing actual world red to humans* to be a necessary property of *red*), then the wavelength of light that produces the color *red* will be a contingent property of *red*. This is because, in the instance of the possible world, the light wavelength that we in the actual world associate with the color *green* appears to humans how *red* appears in the actual world.

On the other hand, if you take a certain wavelength of the light spectrum to be a necessary property of *red*—or surface properties of an object’s ability to reflect a certain wavelength of light, or the human physical ability to pick up on and process a certain wavelength of light, etc.—then this physical property, regardless of how it appears to humans, will be *red*. Therefore, a contingent property of red would be *appearing actual world red to humans*. Thus, even though in our possible world *red* appears to humans like *green* appears in the actual world, this color would regardless be *red* given its physical properties.

With this in mind, we see there is no identity relation between the physical properties of color and its appearance to humans: we can distinguish between possible inverted world *red* and actual world *red* using references to the contingent properties that *red* possesses on each world, depending on what we took for *red*’s necessary properties. If we took the physical properties of *red* to be necessary then we can discuss the differences in how *red* appears in each world, and if we took the appearances properties of *red* as necessary then we can discuss the differences in how *red* is physically grounded in each world. Thus—in the same fashion that Kripke argued that *pain* could not be *C-fibers firing*—how a color appears to humans cannot be a physical property, i.e., it cannot be a brain state or an intentional object. Therefore, qualia exist.

This argument, if it has merit, also undercuts the notion that the content of experience is external and, by extension, that color must necessarily be a property of external objects.

Indeed, we can easily imagine color—as it is experienced by humans—as contingent to objects, considering that the human visual spectrum could have evolved to be somewhere else on the light spectrum. For example, imagine a possible world where human color sight developed to have similar capabilities to those of domestic canines. Now imagine how a red ball from the actual world would appear to humans in this canine color vision possible world. It would appear to us in this possible world as it appears to canines in the actual world, that is, it would appear *grey*. Now, the question becomes, does this ball possess the property of being colored *red* or *grey*? Externalists would have to affirm both, since both properties would have to simultaneously exist in the external object, regardless of the human experience of it, if externalism was true. An eliminativist may counter that the ball possesses the property of being colored red in the actual world but the property of being grey in the canine vision world. But this muddles the fact that this ball has not changed any of its objective, physical properties between worlds. It is still the same numerically identical ball, and therefore should not be able to possess different objective properties depending on what world it is in.

I imagine some eliminativists will just flatly deny that what I called the property of *appearing x to humans* exists. All that exists, they will say, are objective, physical, quantifiable properties, and any seeming qualitative property is just an illusion. They may claim that “color properties” reside in the physical functions of the human eye and brain, and if human beings evolved to have different color experiences, then these physical functions would have all evolved differently as well.<sup>66</sup> I think such a response is plausible on Churchland’s view, who holds that phenomenal experience really just is the experience of physical brain states.<sup>67</sup> However, it would still need to be explained how seemingly

---

<sup>66</sup> See Dennett, *Consciousness Explained*, 389-398.

<sup>67</sup> Churchland, *A Neurocomputational Perspective*, 54-55.

phenomenal experience leads to true justified belief about the *external* world (which Dennett affirms we are able to gain on eliminativism)<sup>68</sup> if the way the external world is perceived is based upon contingently developed physical functions.

Someone may counter that, considering the argument developed in this section, true justified belief about the external world is impossible to gain, since sense properties are only contingently related to physical properties. There are several ways this objection can be answered. The first is to admit that, in the same vein as Edmond Wright, mental images do not match reality at all: objective reality lacks color and the kind of depth that we experience it as having.<sup>69</sup> This, however, is an undesirable concession, since it is admitting that knowledge of the external world really can't be gained. The second is to maintain that human beings evolved in such a way as to experience the external world as it actually is. However, this position, although plausible, would be difficult to prove since we could never get outside ourselves, so to speak, to check if our experiences were accurate to reality. The third is to give up on naturalism altogether.<sup>70</sup>

In any case, given the contingent relation between the physical and experiential properties of color that naturalism entails, it seems plausible that qualia do exist. This, along with the demonstration of the reasonableness of qualia in section 3.3 of this chapter, should motivate us to adopt qualia realism without worry. The existence of such phenomenal properties will be key to the development of my theory of moral perception in part II, but before that, there is a more immediate concern that the existence of qualia presents for us, namely, the mind-body interaction problem.

---

<sup>68</sup> Dennett, *Consciousness Explained*, 364.

<sup>69</sup> Edmond Wright, "Why Transparency is Unethical," in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 343.

<sup>70</sup> Since I am not defending naturalism, but only assuming it for the purposes of this argument, I do not have a particular answer for this objection.

## 4. Mind and Body Interaction

### 4.1 Introduction

The purpose of this chapter is to develop a reasonable metaphysic of the human substance, that is, in terms of mind-body interaction. It begins with commitment to phenomenal, non-reducible consciousness, contrary to any eliminativist theories. In the previous chapter, I justified belief in a distinctly mental realm through a defense of phenomenal consciousness, arguing on the basis of taking seriously the first-personal experience of seemingly qualitative experiences. If qualia exist—or even *seem* to exist<sup>1</sup>—then the mental is constituted, at least in part, by non-reducible, purely immaterial phenomenal content, counter to eliminativism. Therefore, I will not consider reductivist physicalist theories further.

This chapter also begins by assuming causal closure, i.e., the notion that effects we observe in nature have causes—and cannot fail to have causes—constituted of the same kind of stuff<sup>2</sup> as those effects we observe, e.g., physical effects must have physical causes. I have previously argued in chapter 1 that leaving room for indeterminacy between causes and effects is unreflective of how we understand causal events and develop causal explanations—hence, effects have definite causes, even if those causes are unknown. Moreover, denial of causal closure leaves us with the infamous mind-body interaction problem—that is, the problem of how two different kinds of stuff (e.g., mental stuff and physical stuff) can be involved in the same causal chain. By affirming causal closure, this chapter begins with an

---

<sup>1</sup> The *seeming* nature of qualia is itself qualitative and phenomenological. See also Galen Strawson, *Mental Reality*, 2<sup>nd</sup> ed. (Cambridge, MA: MIT Press, 2009), 51-52.

<sup>2</sup> I am using *stuff* here in a technical sense, as referring to a way of fundamentally being, i.e., being physical, or mental, or what have you. For clarity's sake, I am refraining to utilize the more popular word *substance* in this context as I will be using *substance* to refer to Aristotelian substances later in the chapter, although I will still be making reference to *substance dualism* for purposes of discussing dualist theories of mind.

extremely skeptical view of substance dualism. The *prima facie* available options we have as theories of mind-body interaction therefore are epiphenomenalism, non-reductive physicalism, and panpsychism. However, since epiphenomenalism is undesirable as a theory of mental causation (since the mental in this case fails ultimately to be efficacious), it will not be considered.

It will be shown that those theories which allow for the intrinsic nature of causal powers (i.e., the existence of *real* powers) in the human substance are best suited to explain the causal efficacy of human beings while maintaining a commitment to irreducible mental properties, although causal closure remains a problem. I will begin from the physicalist's position with a discussion of non-reductive physicalism. I will then move to the panpsychist's argument in denial of a distinctly physical realm. Following this, I will provide reasons for rejecting the existence of a physical realm altogether, that is, for affirming idealism. This will include responding to objections to idealism at length, since idealism is *prima facie* an unattractive theory. Nevertheless, I conclude that any view which affirms irreducible mental properties and allows for the intrinsic nature of human causal powers—panpsychism, idealism, and substance dualism—is a reasonable metaphysic of the human substance.

## **4.2 Non-Reductive Physicalism**

Needless to say, non-reductive physicalist theories are *physicalist* theories. However, there is some ambiguity as to the definition of *physicalism* and *physical*, and to what kind of physicalist ontology any given physicalist theory may be adhering to. I will be taking a general approach to the definition of physicalism, in order to capture the physicalist commitment at its broadest. Namely, I will be defining physicalism as a commitment to the thesis that all that exists is/are—or is instantiated by, grounded in, constituted by, or in some fashion dependent on—those fundamental entities as theorized by natural science. Thus, we

are now unencumbered in terms of finding a definition of *physical* while allowing for disagreement amongst physicalists to persist.<sup>3</sup>

Non-reductive physicalist theories, then, are those theories that affirm the reality of non-reducible mental properties and entities but nevertheless claim these properties and entities depend on physical properties (i.e., properties affirmed by natural science) for their instantiation. Hence, I am defining non-reductive physicalism in terms of Jaegwon Kim's own definition of a minimal physicalist theory (i.e., a theory that exhibits the minimal commitments necessary to be considered a *physicalist* theory), with the relation between a physical property and its instantiated mental property as one of *supervenience*.<sup>4</sup> According to Kim, supervenience is a relation between property kinds and not instances, thus allowing for multirealizability, i.e., the ability of mental properties to be realized by multiple physical properties.<sup>5</sup> Moreover, Kim points out that non-reductive physicalist theories, in their quest to solve the mind-body problem, carry presumptive commitment the causal efficacy of non-reducible mental content, and, second, to causal closure—commitments that are affirmed in the present analysis as well.<sup>6</sup>

However, as an undesirable consequence, these two commitments can lead to causal overdetermination in a physicalist ontology, as demonstrated by Kim.<sup>7</sup> The reason why non-reductive physicalism results in overdetermination is because, due to supervenience, the

---

<sup>3</sup> See also Daniel Stoljar, *Physicalism* (New York: Routledge, 2010), 69-95, for a discussion of the difficulties involved in defining *physicalism* in terms of scientific theorizing. As Stoljar points out, a concept of physicalism based on our notion of intuitively physical objects is too restrictive, while a concept of physicalism based on what science theorizes is too broad, allowing to the existence of objects that are intuitively nonphysical. With my definition here, I attempt to capture physicalist commitments to science and naturalism without restricting the nature of what can possibly exist if naturalism is true.

<sup>4</sup> Jaegwon Kim, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation* (Cambridge, MA: MIT Press, 1998), 6, 15.

<sup>5</sup> *Ibid.*, 6, 9-10, 19-26.

<sup>6</sup> *Ibid.*, 37-38.

<sup>7</sup> *Ibid.*, 43-44.

instantiation of every mental property depends on the existence of some physical property. As Kim explains, the way to account for how a mental event has causal bearing on a physical event via the supervenience relation is to say that mental event  $m$ , occurring at time  $t$ , causes physical event  $p$ , and further that the causal relation that holds between  $m$  and  $p$  is a result of  $m$  being an event of mental kind  $M$  and  $p$  being an event of physical kind  $P$  (i.e., mental event  $m$  is of the kind  $M$  that causes physical kind  $P$ ). However, this account would be a violation of causal closure, since  $p$  would not have a physical cause. And if we said that  $p$  did have a physical cause,  $p^*$ , then  $m$  having causal bearing on the production of  $p$  would be case of overdetermination. Further, any mental event  $m^*$  that caused  $m$  must itself have had a physical realizer,  $p^*$ , due to the commitment to physicalism.  $p^*$  would therefore be the cause of  $p$ , and  $m$  could have no causal role to play.<sup>8</sup> Kim concludes that, if one is committed to both physical causal closure and to phenomenal realism, then the only available mind-body interaction option seems to be epiphenomenalism.<sup>9</sup>

There is also a problem with the notion of supervenience as an account of phenomenal consciousness, as is evident from Chalmers' analysis of the supervenience relation. Chalmers begins his analysis by making a distinction between logical supervenience and natural supervenience, the former being a relation of logical entailment, and the latter being one of nomological correlation.<sup>10</sup> He then argues that, since it is logically possible that phenomenal consciousness does not exist, it can only be naturally supervenient if it is supervenient at all.<sup>11</sup> He demonstrates this logical possibility through his famous conception of a phenomenal zombie: a physical and functional silicon isomorph of himself with no

---

<sup>8</sup> Ibid., 41-46.

<sup>9</sup> Ibid., 118-120.

<sup>10</sup> David J. Chalmers, *Conscious Mind: In Search of a Fundamental Theory* (Oxford: Oxford University Press, 1996), 34-36.

<sup>11</sup> Ibid., 37-38.

experiential mental content.<sup>12</sup> If it is conceivable that such a zombie can exist, then the supervenience relation cannot be one of logical entailment.<sup>13</sup> Further, Chalmers argues that if consciousness was logically supervenient, it should be able to be reductively explained in terms of functional roles. The whole point of the non-reductive physicalist project however is that consciousness is not reducible in this manner, and therefore consciousness must be naturally supervenient if supervenient at all.<sup>14</sup>

Chalmers further makes the point that, given the natural laws of the actual world, there are more constraints on what can be considered naturally possible than there are on what can be considered logically possible.<sup>15</sup> Through this observation, he highlights a connection between natural possibility and natural law-like correlations. Given Chalmers' analysis here, what we observe of nomological relations should inform us as to what is naturally possible. However, what we consider naturally possible is supposed to inform us of what nomological relations exist. In other words, analyzing what is naturally possible should shed light on whether supervenience exists, but it is only by observing supervenience—and thus knowing that it exists—that we would know it is naturally possible. This epistemic circularity seems to point to the impossibility of definitively demonstrating that the supervenience relation is possible. In fact, given this circularity, nothing can be demonstrated as possible unless it is demonstrated as actual. This problem calls into question the use of supervenience *as an explanation* for the existence of phenomenal consciousness. Something more needs to be said than merely claiming a nomological connection; this law-like

---

<sup>12</sup> Ibid., 94-95.

<sup>13</sup> Ibid., 97.

<sup>14</sup> Ibid., 44-50.

<sup>15</sup> Ibid., 36-37.

connection itself is in need of explanation. Without said explanation, non-reductive physicalism seems untenable as an account of non-reducible phenomenal consciousness.

Fundamentally, the problems that non-reductive physicalism faces aren't due to its commitment to causal closure or the non-reducibility of the mental, but to its commitment to a certain understanding of what constitutes the physical. Importantly, non-reductive physicalists hold that the physical is best understood as an ontologically distinct kind of stuff from the mental. Our next theory of mind, panpsychism, rejects this understanding of what constitutes the physical.

### 4.3 Panpsychism

Like substance dualists and non-reductive physicalists, panpsychists take our first-personal phenomenal experiences of our conscious states as epistemologically authoritative: because we have these phenomenal experiences, we know these phenomena exist.<sup>16</sup> However, unlike substance dualists and non-reductive physicalists, panpsychists attempt to give an account of mental phenomena that is embedded in the fundamental nature of physical reality in a way that renders these phenomena neither wholly distinct nor emergent,<sup>17</sup> thereby allowing for them to interact in a causal chain with physical entities. Hence, panpsychism is the view that mental content is real and non-reducible because it pervasively occurs in nature (as opposed to merely arising when matter reaches a certain complexity or obtains a certain structure) even at the most fundamental level. Chalmers himself defines panpsychism as “the thesis that everything is conscious, or at least that fundamental physical entities are conscious.”<sup>18</sup> Strawson himself claims that “[b]eing [i.e., existence] is essentially experience-

---

<sup>16</sup> Galen Strawson, “Mind and Being: The Primacy of Panpsychism,” in *Panpsychism: Contemporary Perspectives*, ed. Godehard Bruntrup and Ludwig Jaskolla (Oxford: Oxford University Press, 2016), 84. See also Chalmers, *Conscious Mind*, xii-xiii.

<sup>17</sup> Strawson, “Mind and Being,” 82.

<sup>18</sup> David J. Chalmers, “Panpsychism and Protopanpsychism,” *The Amherst Lecture in Philosophy* 8 (2013): 1.

involving—where ‘experience’ denotes any sort of conscious experience whatever, including the most primitive forms of ‘mere’ sensation.”<sup>19</sup> It is important to emphasize this metaphysical fundamentality of phenomenal experience since holding to this is what sets panpsychists apart from non-reductive physicalists.

Thus, like non-reductive physicalists, panpsychists retain a commitment to physicalism. However, unlike non-reductive physicalists, panpsychists redefine the physical to include mental phenomena. As Strawson argues, this redefinition is justified since we lack a full concept of the physical (or physical *stuff*, what Strawson calls “matter”); our lacking a complete concept is why mind-body interaction is such a problem.<sup>20</sup> He claims:

Many take this to be the problem of how mental phenomena can be physical phenomena *given what we already know about the nature of the physical*. But those who think this are already lost. For the fact is that we have *no* good reason to think that we know anything about the physical that gives us any reason to find any problem in the idea that mental phenomena are physical phenomena [emphasis Strawson’s].<sup>21</sup>

Moreover, Strawson argues that physicalists “must grant that experiential phenomena are real, concrete phenomena, for nothing in this life is more certain.”<sup>22</sup> Therefore, physicalists should take the position that experiential phenomena are physical phenomena.<sup>23</sup> Hence, for panpsychists, the commitment to the existence of phenomenal consciousness is merely part of the commitment to physicalism in general, since there is no real metaphysical distinction between the mental and the physical to begin with.

---

<sup>19</sup> Strawson, “Mind and Being,” 80.

<sup>20</sup> Strawson, *Mental Reality*, 98-99.

<sup>21</sup> Galen Strawson, “Real Materialism,” in *Real Materialism: and Other Essays* (Oxford: Oxford University Press, 2008), 20.

<sup>22</sup> *Ibid.*, 21. The term that Strawson uses here is *materialists*, not *physicalists*. However, unlike many others, Strawson utilizes the terms *physicalist* and *materialist* interchangeably. See Strawson, “Mind and Being,” 83, where Strawson explicitly states as such.

<sup>23</sup> Strawson, “Real Materialism,” 21.

Also important to recognize is that panpsychism possesses a built-in framework for the existence of intrinsic properties, which for panpsychists would be experiential properties. For example, Strawson holds to what he calls “a form of *pure* panpsychism,” i.e., what he takes “to be the view that experientiality is all there is to the intrinsic nature of concrete reality.”<sup>24</sup> Regarding intrinsic properties, Chalmers speculates the following:

[I]t is often noted that physics characterizes its basic entities only *extrinsically*, in terms of their relations to other entities, which are themselves characterized extrinsically, and so on. The intrinsic nature of physical entities is left aside. Some argue that no such intrinsic properties exist, but then one is left with a world that is pure causal flux (a pure flow of information) with no properties for the causation to relate. If one allows that intrinsic properties exist, a natural speculation given the above is that the intrinsic properties of the physical—the properties that causation ultimately relates—are themselves phenomenal properties.<sup>25</sup>

As Chalmers points out above, the existence of these phenomenal intrinsic properties helps explain the ordered nature of causation and provides a place for real causal powers to exist. Without this, a metaphysic is at a disadvantage since it would have difficulty making room for anything except Humean causation, which was shown to be inadequate in chapter 1. Therefore, panpsychism seems like a good option to adopt as a metaphysic in that it maintains irreducible mental properties and causal closure. Moreover, because panpsychism makes room for intrinsic properties, it makes room for the substance theory of causation: mental properties, as intrinsic properties, could play the role of powers given their ability to affect the behavior of objects.

#### **4.4 Problems with Panpsychism**

Panpsychism is not without its problems, however. One such problem is known as the combination problem. Basically, the combination problem arises when questioning how it may be that the macrolevel phenomenal consciousness of human beings can arise from the

---

<sup>24</sup> Strawson, “Mind and Being,” 81.

<sup>25</sup> David J. Chalmers, “Facing Up to the Problem of Consciousness,” *Journal of Consciousness Studies* 2 (1995): 217.

microlevel consciousnesses of individual fundamental particles<sup>26</sup> (or whatever one's panpsychist theory considers the basic individual instances of phenomenal content).<sup>27</sup> Or as Strawson puts it, how do the many experiences of many subjects come together to make one unified experience of a single subject?<sup>28</sup> Combinatory problems of this type plague all physicalist theories that attempt to account for phenomenal consciousness by an appeal to the unity, coordination, or complexity of constituent units—i.e, any theory that denies the mereological simplicity of conscious beings.<sup>29</sup> Whereas all other physicalist theories have trouble providing an account of how consciousness itself *emerges*, panpsychism has set itself up for the unique problem of providing an account of how *macrolevel* consciousness *emerges*. In that sense, panpsychism seems to not have made much progress over emergentism.

Another problem with panpsychism is one related to the mind-body problem, that is, the problem of explaining how internal phenomena interact with an external world. This problem is much different for panpsychists than it is for non-reductive physicalists or substance dualists, however. For the latter theories, this problem results from taking the mental and the physical to be wholly distinct kinds of stuffs, which results in the difficulty of trying to fit them in the same causal chain. For the panpsychists, on the other hand, this problem arises from maintaining *universal internal phenomenology* through a commitment to the consciousness of fundamental particles as well as *universal external physicality* through a commitment to physicalism, thereby creating a natural tension. This tension comes to the fore

---

<sup>26</sup> David J. Chalmers, "The Combination Problem for Panpsychism," in *Panpsychism: Contemporary Perspectives*, ed. Godehard Bruntrup and Ludwig Jaskolla (Oxford: Oxford University Press, 2016), 179-180.

<sup>27</sup> Strawson, "Mind and Being," 102.

<sup>28</sup> *Ibid.*

<sup>29</sup> David Barnett, "You are Simple," in *The Waning of Materialism*, ed. Robert C. Koons and George Bealer (Oxford: Oxford University Press, 2010), 161.

in explaining how experiential aspects of objects and properties interact with non-experiential aspects, in much the same fashion that non-reductive physicalists and substance dualists have trouble in explaining how mental properties can interact with physical properties. This problem is most pronounced when one type of causal explanation—for example, a non-experiential molecular explanation of chemical changes—exhaustively explains a particular causal occurrence while the other type—the experiential aspects of said same chemical changes—seems to play no causal role. As such, similarly to non-reductive physicalists, panpsychist accounts of causation run the risk of overdetermination if we take both experiential and non-experiential explanations seriously. What is doing the causing in agent causation? Is it the external excitement of neurons or the internal will of a subject? The panpsychist affirms that these are just external and internal aspects of the same thing, but there seems to be a risk of advocating for a form of epiphenomenalism if non-experiential causal accounts are given primacy.

We can see this problem more clearly if we consider a thought experiment. Let's say that I pick up a rock. In this scenario, given panpsychism, there's me, a sophisticated macrolevel conscious being, and a rock, a probably not-as-sophisticated conscious being. I internally perceive the rock, internally decide to pick it up, and internally exert my will: these are the experiential aspects of this causal occurrence. Externally, my body moves and the rock gets picked up: this is the non-experiential, third-personal aspects. What is the relationship between the experiential and the non-experiential? If we explain this relationship by claiming that there is no metaphysical difference—the relationship is one of identity—then we run the risk of saying everything is experiential and nothing is non-experiential. Thus, with a commitment to the experiential and without an account of the experiential's relationship to the non-experiential, we risk idealism. Perhaps this isn't really a big problem for the panpsychist—Strawson himself acknowledges the movement towards idealism his

theory suggests.<sup>30</sup> But there is a sense, even for Strawson, that panpsychism is not idealism, at least a form of idealism that holds the internality of all existents.<sup>31</sup>

Since panpsychism still succumbs to the problems inherent in attempting to bridge the gap between an external and internal world, it falters in the face of the mind-body problem. Perhaps where panpsychism fails is where physicalism in general fails—not with positing universally pervasive consciousness so much as trying to retain the notion of external, mind-independent, material stuff. The solution to the problems of panpsychism therefore might be to collapse the external to the internal and assume the mentality of all existents.

What if we were to assume the mentality of all existents? That is, why should we not posit idealism? We would still be able to hold to the two commitments that began this discussion—the reality of irreducible phenomenal consciousness and causal closure (that is, mental activity having mental causes and mental effects as opposed to physical causes and effects)—as well as avoid the problems inherent in accepting a distinctly physical, external reality. The idea of idealism is objectionable to many, however, and further it does not seem to cohere with our natural intuitions concerning what we perceive as the mind-independent world around us. Much more needs to be said if we are going to take idealism as a serious option.

#### **4.5 Idealism**

Berkeley, in *A Treatise Concerning the Principles of Human Knowledge*, argues against the existence of the physical, as it is traditionally understood (what Berkeley calls “matter”), on the basis that, even if the physical did exist, we could not know it exists. Berkeley argues that since matter is “passive and inert,” it “cannot be an agent or efficient

---

<sup>30</sup> Strawson, “Real Materialism,” 23, 49.

<sup>31</sup> Strawson, “Mind and Being,” 93-94.

cause.”<sup>32</sup> Moreover, since what is perceivable are qualitative properties, matter itself is “unperceivable, as being devoid of all sensible qualities, and so cannot be the occasion of our perceptions...as when the burning my finger is said to be the occasion of the pain that attends it.”<sup>33</sup> Therefore, in order to avoid the skepticism that arises out of taking matter to be distinct from mind, Berkeley adopts idealism. Or, in his own words:

So long as we attribute a real existence to unthinking things, distinct from their being perceived, it is not only impossible for us to know with evidence the nature of any real unthinking being, but even that it exists. Hence it is that we see philosophers distrust their sense, and doubt of the existence of heaven and earth, of everything they see or feel, even of their own bodies. And, after all their labour and struggle of thought, they are forced to own we cannot attain to any self-evident or demonstrative knowledge of the existence of sensible things...I can as well doubt of my own being as of the being of those things which I actually perceive by sense; it being a manifest contradiction that any sensible object should be immediately perceived by sight or touch, and at the same time have no existence in nature, since the very existence of an unthinking being consists in being perceived.<sup>34</sup>

In this way, Berkeley turns to idealism in order to maintain the metaphysical reality of what we perceive and, moreover, our ability to gain knowledge of “sensible things,” i.e., the objects of perception.

Robert Adams argues for idealism along similar lines, stating that the difficulty of discovering what objects are like within themselves naturally leads one to believe that perhaps there is nothing to what objects are like within themselves.<sup>35</sup> To understand this reasoning, recall the argument in chapter 3 that humans could have evolved to see leafy green plants phenomenally as a different color. This is the case with any sensible thing: it could

---

<sup>32</sup> George Berkeley, *A Treatise Concerning the Principles of Human Knowledge*, §69. Note I am using section numbers, found in the text of the *A Treatise Concerning the Principles of Human Knowledge*, in order to reference *A Treatise Concerning the Principles of Human Knowledge*, while I will be using the page numbers from the Roger Woolhouse edited volume of Berkeley’s works in order to reference *Three Dialogues Between Hylas and Philonous* (see footnote 49).

<sup>33</sup> Ibid.

<sup>34</sup> Ibid., §88.

<sup>35</sup> Robert M. Adams, “Idealism Vindicated,” in *Persons: Human and Divine*, ed. Peter van Inwagen and Dean Zimmerman (Oxford: Oxford University Press, 2007), 36.

have appeared differently to human perception. This relation between perception and the objects of perception calls into question our ability to understand what things are like in themselves: perhaps we have little to no understanding of mind-independent reality.<sup>36</sup> Moreover, if we allow that at least some properties that we perceive do not belong to mind-independent objects (such as color properties, or taste and smell properties), then it is perhaps possible (or at least, conceivable) that no properties do.<sup>37</sup> All properties (or at least a vast majority of properties) exist internally in the mind of subjects.<sup>38</sup>

Adams sees this idealistic thinking as a reaction to modern philosophers' mechanistic model of the universe, which, according to Adams, is understood "in terms of geometrical properties and motions of bodies which interact only by touching and pushing each other."<sup>39</sup> Given this modern conception of interaction, it is natural to assume that all causes and effects have mechanistic explanations, and the properties and behavior of objects are "geometrical."<sup>40</sup> According to Adams, however, this model leaves something lacking. In response to Descartes' assertion that corporeality is merely extension, Adam argues that "a system of spatiotemporal relationships constituted by sizes, shapes, positions, and changes thereof is too incomplete, too hollow, as it were, to constitute an ultimately real thing or substance."<sup>41</sup> Contending that this conception of corporeality is inadequate, he advocates for

---

<sup>36</sup> See also Edmond Wright, "Why Transparency is Unethical," in *The Case for Qualia*, ed. Edmond Wright (Cambridge, MA: MIT Press, 2008), 342-345, where Wright talks about the lack of resemblance between qualia and the objects that gave rise to said qualia.

<sup>37</sup> Todd Buras and Trent Dougherty, "Parrying Parity: A Reply to a Reidian Critique of Idealism," in *Idealism: New Essays in Metaphysics*, ed. Tyrone Goldschmidt and Kenneth L. Pearce (Oxford: Oxford University Press, 2017), 2-3.

<sup>38</sup> See Adams, "Idealism Vindicated," 47. Adams specifically defends a version of idealism where some causal relations are external. This is his way of avoiding solipsism.

<sup>39</sup> Adams, "Idealism Vindicated," 38.

<sup>40</sup> *Ibid.*

<sup>41</sup> *Ibid.*, 40.

a richer conception to fill into the formal and structural ontology he considers the modern mechanistic ontology to be.<sup>42</sup> Adams then goes on to question whether or not it is secondary qualities that fill in this gap:

Within the visual field the colors literally fill in the shapes; and it is because shapes need a filling that we can hardly imagine, visually, a shape without some chromatic property. And it is because of the qualitiveness of colors that they bring to the context something that is not merely formal and structural. In a more general way, then, we may conjecture that the reality of a substance must include something intrinsic and *qualitative* over and above any formal or structural features it may possess.<sup>43</sup>

Adams then affirms “first, that substances must have intrinsic non-formal qualities, and second, that qualities of consciousness, or qualities very like them, are the only intrinsic non-formal qualities of substances.”<sup>44</sup> Therefore, it seems, conscious properties must belong to all substances.

One may recognize that this conclusion—that all intrinsic properties are conscious properties, and that conscious properties are universally possessed by all entities—is both compatible with idealism and panpsychism. Adams admits as much.<sup>45</sup> In order to distinguish his position, Adams maintains that, under idealism, spatiotemporal relations reduce to internal mental content, which Adams believes is idealism’s differentiating feature.<sup>46</sup> We can see now how my above objection to panpsychism—that it fails under something like the mind-body problem—does not work against idealism. In the above objection, I called into question how the internal exercise of my will was able to cause effects to my external physicality and the rock’s external physicality. However, if spatiotemporal relations—

---

<sup>42</sup> Ibid.

<sup>43</sup> Ibid.

<sup>44</sup> Ibid., 47.

<sup>45</sup> Ibid.

<sup>46</sup> Ibid.

relations that give corporeality to my body and the rock's body—were reduced to mental content, then nothing external (or nothing much external) is happening when I move the rock.

#### 4.6 Objections to Idealism

In *The Nature of Perception*, John Foster—similarly to Berkeley and Adams—defends the claim that, in order for humans to have perceptual access to the world, the world must be both logically dependent on the human mind and fundamentally consist in human mental facts.<sup>47</sup> According to Foster, these two claims are what constitute idealism.<sup>48</sup> However, if the world is dependent on and consists in the facts of immediate mental states, what it seems we are left with is solipsism: for indeed, if I collapse the perceived world to my perceptions of the world itself, and deny the existence of anything beyond those perceptions, then other minds except my own do not exist. Shouldn't we reject idealism based on this bent towards solipsism?

One solution to this problem is to embrace the conclusion that there is only one mind while giving an account of a quasi-mind-independence. This is what Bernardo Kastrup does when he argues that there exists only one universal consciousness that each individual consciousness derives its being from through a process of mental dissociation, i.e., the psychiatric concept that describes the discontinuity of some mental content with other mental content in a single mind. For Kastrup, each human being is a bit of universal consciousness that is logically disconnected (as opposed to metaphysically disconnected) from the universal consciousness as a whole. In this fashion, Kastrup is able to maintain some semblance of mental privacy.<sup>49</sup>

---

<sup>47</sup> John Foster, *The Nature of Perception* (Oxford: Oxford University Press, 2000), 249.

<sup>48</sup> *Ibid.*, 255.

<sup>49</sup> Bernardo Kastrup, "On the Plausibility of Idealism: Refuting Criticisms," *Disputatio* 9 (2017): 16-18.

The other, and, in my opinion, more attractive solution to the problem of solipsism is to reject Foster's denial of mind-independence as inherent to idealism. For one, we are naturally inclined to believe that, even if idealism is true (and solipsism is false), there is at least some mind-independence: if more than one mind exists, each mind is independent of all others. Foster himself even admits that idealism, if it is to be non-solipsistic, requires an external basis.<sup>50</sup> This external basis would both assume and allow for mind-independence.

What would the character of this external basis be? The most straightforward answer—especially for a theist—would be Berkeley's view or something similar to it.<sup>51</sup> Berkeley held that the existence of a perceived object did not subsist in the perceptions of a human observer but in the perceptions of a divine observer. Thus, objects of our perceptions have mind-independent existence from our minds since they exist in God's mind. In fact, for Berkeley, existence seems to consist solely of being perceived by God and nothing can exist unless it is so perceived.<sup>52</sup> However, as Daniel Flage also clarifies, this doesn't mean that Berkeley holds that God has sensations as humans have sensations: God having sensations would assume God's passibility<sup>53</sup> and Berkeley explicitly denies God is passible.<sup>54</sup> Nevertheless, that God would have perception of his creation is without controversy, as Berkeley points out.<sup>55</sup> The postulation of a transcendent observer undergirding the existence of objects of human perception allows for the existence of a reality independent from human

---

<sup>50</sup> Foster, *The Nature of Perception*, 281-282.

<sup>51</sup> See also Foster, *The Nature of Perception*, 283, where Foster mentions Berkeley's theory as a solution to this problem.

<sup>52</sup> George Berkeley, *Three Dialogues Between Hylas and Philonous*, 160-162. Page numbers taken from combined volume *A Treatise Concerning the Principles of Human Knowledge and Three Dialogues Between Hylas and Philonous*, ed. Roger Woolhouse (London: Penguin Books Ltd, 1988).

<sup>53</sup> Daniel E. Flage, "Berkeley's Archetypes," *Hermathena* 171 (2001): 14.

<sup>54</sup> Berkeley, *Three Dialogues Between Hylas and Philonous*, 162.

<sup>55</sup> *Ibid.*, 161.

minds.<sup>56</sup> In this way, idealism avoids solipsism. Unfortunately, this solution is unattractive to non-theists.

Another objection to idealism is that, if truth conditions are given by phenomenal content, then there can be no such thing as objective truth. Indeed, if reality consists in perceptions, then illusions, dreams, and hallucinations are no less real and true than any other perception. We need physical reality in order to ground truth. As an answer, Kastrup argues that the reason why we come to know that we are wrong about misperceptions, such as with illusions, is because of subsequent perceptions that inform our judgments.<sup>57</sup> Moreover, holding to physicalism presents a bigger epistemological problem:

Lucid contemplation of [idealism and physicalism] shows that idealism attempts to reduce an explanatory abstraction (physically objective matter) to that which articulates and hosts the abstraction in the first place (mind). This is *prima facie* eminently reasonable. Mainstream physicalism, in turn, attempts to reduce mind to mind's own explanatory abstractions, an obvious paradox that constitutes the crux of the "hard problem."<sup>58</sup>

As previously discussed, and as Berkeley argued,<sup>59</sup> holding the physical as ontologically distinct from the mental makes it impossible to get out into the world—to know that what we perceive of the external world is what the world is like in and of itself. If we take the mental to be distinct from the physical, it is hard to make sense of the physical, considering all we know about it is what we perceive in the mental. It seems as if any worldview is a barrier to truth, it's one that admits material objects.

Berkeley acknowledged that the perceptions of a single object as had by different subjects can all be different, but the true reality of an object is grounded in object's

---

<sup>56</sup> Ibid., 182.

<sup>57</sup> Bernardo Kastrup, "Conflating Abstraction with Empirical Observation: The False Mind-Matter Dichotomy," *Constructivist Foundation* 13 (2018): 344.

<sup>58</sup> Ibid., 346.

<sup>59</sup> Berkeley, *Three Dialogues Between Hylas and Philonous*, 176-177.

archetype, which exists in the mind of God.<sup>60</sup> As Flage points out, these archetypes ground identity claims in Berkeley's ontology.<sup>61</sup> Therefore, there is a truth of the matter for each claim about an object of perception. Furthermore, for Berkeley, all qualitative properties of an object really exist in the object—as opposed merely existing in internal qualia separate from the external object itself—and thus our perceptions of an object gives us real immediate knowledge of that object. The example Berkeley uses is that of a tulip: because the tulip consists in its idea with all its qualitative properties, the color property of the tulip truly exists in the tulip itself.<sup>62</sup> Moreover, objects consisting of being perceived does not make them any less real for Berkeley. For Berkeley, to be an idea in this fashion is to have real existence.<sup>63</sup>

But what about dreams and hallucinations? How does Berkeley distinguish what is “real” perceptions from these false ones, if everything is in the mind? Berkeley contends that false perceptions are “faint and indistinct,” have the character of being “dim, irregular, and confused,” and are wholly dependent on the human will. On the other hand, perceptions of the real “are more vivid and clear” and are outside one's control.<sup>64</sup> Counter to Berkeley's assumption, I am uncertain that a good case can be made in support of the notion that real perceptions are wholly outside one's control (or that false perceptions are totally within one's control, for that matter). Nevertheless, it seems pretty evident that the experience of dreams and hallucinations is less orderly and coherent, especially in terms of each instance of a dream or hallucinatory experience logically connecting to other experiences, both real and

---

<sup>60</sup> Ibid., 192-194.

<sup>61</sup> Flage, “Berkeley's Archetypes,” 12.

<sup>62</sup> Berkeley, *Three Dialogues Between Hylas and Philonous*, 144.

<sup>63</sup> Ibid., 176-178. See also Adams, “Idealism Vindicated,” 47, where Adams makes a similar point about Berkeley's idealism.

<sup>64</sup> Berkeley, *Three Dialogues Between Hylas and Philonous*, 182.

false.<sup>65</sup> Because of this, we as perceptual beings are generally able to distinguish between reality and fantasy. Since we have this ability, it cannot be said that an idealist ontology hinders knowledge of objective reality.

Another objection to idealism is that idea that, if idealism is true, then there's no such thing as real causal interaction between objects or substances. For indeed, if idealism is adopted, then one primary reason for adopting it—that it maintains causal closure—becomes moot. Moreover, idealism can't allow for an ontology of substances with causal power and thus would not work with the theory of causation advocated in chapter 1.

In fact, with idealism the existence of the human substance itself is called into question. Berkeley infers the existence of the self, the human substance, from his own inner reflection of himself; however, for Berkeley, the human substance seems to be a *spiritual* substance, which he considers distinct from thoughts and perceptions.<sup>66</sup> It seems as if Berkeley missteps to say he can infer the existence of his own spiritual substance without direct perception given he doesn't allow this kind of inference to the knowledge of material substances.<sup>67</sup> Berkeley, however, addresses such an objection, stating that he refutes the existence of material substances because of the inconsistency involved in the idea of the material, not just on the basis of the material being imperceptible. Nevertheless, it is difficult to see how postulating that a spirit has thoughts and perceptions is anything over and above postulating a mind inhabiting a body, for how do thoughts relate to a spirit? It seems as if we must either say that a spirit just is a being of pure mental states—and nothing over and above those mental states—or we arrive back at the mind-body problem again. In either case, how idealism can account for the existence of the human substance is unforthcoming.

---

<sup>65</sup> See also Adams, "Idealism Vindicated," 48, where Adams discusses this point.

<sup>66</sup> Berkeley, *Three Dialogues Between Hylas and Philonous*, 179-181.

<sup>67</sup> *Ibid.*, 137, 146-149.

Furthermore, Berkeley contends that God gives order to nature, which exhibits lawlikeness.<sup>68</sup> But it is difficult to see how this is anything above occasionalism given that God's mind is fully in charge of the causal order. As Adams demonstrates, idealism *prima facie* doesn't seem to support the idea that substances have any real causal interaction with each other; it seems to be solely an occasionalist theory due to God's causal sustaining role. Adams explains that this is the case because, under a truly interactionist version of idealism, substances would lack the perceptual ability needed to constitute physical reality as we observe it:

[I]ntuitions regarding the reality of physical causation seem better respected in postulating an interactionist causal structure, in which many perceiving substances, corresponding in some way to physical objects, exercise metaphysically real causal influence on each other... And the obvious reason *not* to prefer such an interactionist version of idealism to the occasionalist version is that it seems doubtful that the rudimentary perceptions of those many substances contain enough information for the construction of a causal order as rich and well articulated as that of physics. Specifically we may wonder whether the feelings of substances that do not have fairly advanced geometrical perception can contain enough information for the construction of space and time from intramental resources which is required for an idealist as distinct from a merely panpsychist theory.<sup>69</sup>

Adams states that what could be postulated to support real causation by the substances themselves is “that spatiotemporal relational properties are indirectly tied to qualities of consciousness (or quasi-consciousness) by *belonging to the same substance*;<sup>70</sup> but this fix results in panpsychism since now spatiotemporal relations are not reductive—they exist externally and physically (or in other words, all existents are not completely phenomenal as idealism would hold, but bear some physicality).<sup>71</sup> Adams concludes that a “either a broadly occasionalist idealism or an interactionist panpsychism can account for the causal structure of

---

<sup>68</sup> Ibid., 178.

<sup>69</sup> Adams, “Idealism Vindicated,” 52-53.

<sup>70</sup> Ibid., 53.

<sup>71</sup> Ibid.

physics as well as a typical physicalist view can.”<sup>72</sup> What is left in doubt, however, is the possibility of a truly interactionist version of idealism that is compatible with the affirmation of substance causation.

This is a real problem for idealists if they cannot come up with an answer. For truly, we want to say that there are real properties belonging to a substance that allow it to act and be acted upon (indeed, this chapter concerns providing an account of mental causation in a substance causation ontology). Perhaps the way forward is to think of these causal properties as belonging to the object (or substance) of perception as any perceived properties would, such as color properties. Moreover, the perceptions of the objects involved in causal interaction are not the grounding perceptions of these powers, counter to Adams’ postulation. It could be argued that the existence of these properties, as well as their interactions, are grounded in God’s mind, just as individual substances themselves are grounded in God’s mind. Arguably, how objects exist in God’s mind may be much finer and more detailed than humans can perceive and therefore complex causal powers that we don’t have direct perceptual access to may exist. The question becomes, is this just occasionalism in disguise? In one sense, yes, because without God thinking it nothing would come to be, including the exercise of causal powers. And in another sense, no, because if we affirm Berkeley’s assessment that to be perceived is to be real in a metaphysically serious sense, then causal powers and the causal interactions between substances are very real.

What does this do to our developing a theory of mental causation—where human beings are said to possess real causal powers originating in the exercise of their wills—if we have to admit that human powers, and the exercise thereof, are grounded in the mind of God? If we accept idealism, shouldn’t we have to say that humans lack the capacity for free will? For a possible answer, we can look at philosophical work surrounding divine providence that

---

<sup>72</sup> Ibid., 54.

deals with exactly this kind of problem, given that the existence of a God, as classically conceived, poses an inherent problem concerning human freedom regardless of what ontology one adopts.<sup>73</sup> It should be concluded that this is not a problem unique to idealism and thus shouldn't be treated as such.

Another objection to idealism goes like this: if the world is exactly as humans perceive it, then how can different species have different perceptual experiences of the same object? For example, a human may see a ball as red that a dog sees as grey. If the ball merely is the perception of it, then how can it consist in multiple perceptions? I've discussed a problem similar to this in the last chapter concerning how externalists about experience might situate color properties that appear to two different species two different ways. I concluded that, because we cannot externally situate such contradictory properties, we had to reject externalism. Now we are considering an internalist theory of experience; however, the properties in question are still externally grounded in a divine mind as opposed to in an outer physical world. Therefore, the problem is still similar. What property belongs to the ball, now internally conceived? Red or grey?

This problem can be met similarly to how we met the previous objection concerning the grounding of objective reality. Since the world that human beings find themselves in is not constituted by human perceptual content, but by divine mental content, the world is not exactly as human beings perceive it anyway since human beings have perceptual limitations. Moreover, just as human beings can have different perceptions of the same object, so it can be inferred that different species can also have different perceptions of the same object. What possesses the properties of the ball in question aren't any species' ideas or perceptions of that

---

<sup>73</sup> See David S. Oderberg, "Divine Premotion," *International Journal for Philosophy of Religion* 79 (2016): 207-222, and Robert C. Koons, "Dual Agency: A Thomistic Account of Providence and Human Freedom," *Philosophia Christi* 4 (2002): 397-411, where Oderberg and Koons attempt to develop a coherent account of humans exercising causal powers and God causing the exercise of those same powers while human freedom is maintained.

object, but the archetype of the ball in the mind of God.<sup>74</sup> According to Flage’s analysis, an archetype is that to which a perception of an object is held in reference, which grounds all identity claims;<sup>75</sup> moreover, an archetype contains within it all possible perspectives of the object it is the archetype of.<sup>76</sup> Perhaps the archetype of the ball in question contains both dog perspectives and human perspectives (as well as an unknown number of other species-based perspectives). Further, from the dog perspective the property *grey* is perceptible and from the human perspective the property *red* is perceptible. Both *red* and *grey*, however, belong to the archetype.

#### 4.7 Substance Dualism Revisited

The downfall of idealism as presented here is that it would only be appealing to theists—non-theists could do nothing with it since it requires the positing of a divine observer in order to ground objectivity. Thus, anyone without a commitment to the existence of God would find cause to dismiss it outright. Perhaps the best option for the non-theist would be a single universal consciousness, such as argued for by Kastrup. Or, considering the deistic connotations of Kastrup’s theory, perhaps panpsychism would be better fitting for a non-theist committed to naturalism. However, with panpsychism, the problem of bridging the gap between the internal and external has yet to be solved, and problems concerning giving an account of mental causation still remain.

What if we turned back and considered substance dualism? We initially rejected dualism on the basis of its closure violations. Yet now we have cast doubt on panpsychism and idealism because of problems inherent in those theories, ones that do not seem to give them an advantage over substance dualism. What if we could explain mental efficacy under a

---

<sup>74</sup> Berkeley, *Three Dialogues Between Hylas and Philonous*, 193-194.

<sup>75</sup> Flage, “Berkeley’s Archetypes,” 7-9.

<sup>76</sup> *Ibid.*, 11.

dualist ontology that doesn't violate closure, one that gives an account of mind-body interaction? There have been recent attempts by dualists to do just this via appeals to mutual manifestation.

Recall from chapter 1 that mutual manifestation is the thesis that a given manifestation event is the product of two or more powers or dispositions manifesting themselves simultaneously. As Christopher Austin points out, according to this thesis, “there *is no such thing* as a manifestation of a *single* dispositional property,”<sup>77</sup> which “is in stark contrast to our usual stimulus-response model” of how substance causation works.<sup>78</sup> In chapter 1, I criticized the mutual manifestation thesis for being fundamentally an anti-powers view of causation since according to mutual manifestation, the manifestation of power consists not in the activity of essential monadic properties but in the extrinsic contingent relations between properties. In this chapter, however, I will revisit it in order to explore whether it helps substance dualism solve the mind-body interaction problem. Ultimately, I argue, it does not.

One dualist theory that takes advantage of the mutual manifestation thesis in order to solve the interaction problem is that of Sophie Gibb. On her theory, Gibb tries to make sense of mental events as *double preventers*.<sup>79</sup> A double preventer prevents a certain manifestation event which, if manifested, would prevent a targeted manifestation event from occurring.<sup>80</sup> In other words, a double preventer allows for the occurrence of a targeted or desired manifestation event which would otherwise not manifest. Gibb illustrates the function of a

---

<sup>77</sup> Christopher J. Austin, “Is Dispositional Causation Just Mutual Manifestation?,” *Ratio* 29 (2016): 235-236.

<sup>78</sup> *Ibid.*, 236.

<sup>79</sup> S.C. Gibb, “Mental Causation and Double Prevention,” in *Mental Causation and Ontology*, ed. S.C. Gibb, E. J. Lowe, and R.D. Ingthorsson (Oxford: Oxford University Press, 2013), 194.

<sup>80</sup> *Ibid.*, 199-200.

double preventer through the example of a game at a fair. In this game, the targeted manifestation is the breaking of a glass bottle by way of a ball thrown by a player. Unknown to the player, however, an invisible barrier would prevent the ball from reaching the bottle if the ball was thrown. Nonetheless, just as a player throws the ball, a fair worker removes the barrier by pressing a button—a *double preventer*—resulting in the targeted effect occurring—the breaking of the glass bottle.<sup>81</sup>

Gibb maintains that “double preventer events are not causes of the event that they have prevented from being prevented.”<sup>82</sup> Thus, a mental event that acts as a *double preventer* does not violate causal closure. However, this doesn’t mitigate the double preventer’s necessity for the effect’s occurrence, since without the double preventer permitting the event, the event would not have occurred. Therefore, as Gibb claims, “the role of an event that permits a cause to bring about an effect is no less important than the role of the cause.”<sup>83</sup> Moreover, she argues that the double preventer’s role in a causal event is “an objective one, not merely an explanatory one.”<sup>84</sup> Thus, by way of this causal importance of a double preventer, Gibb attempts to account for the causal efficacy of the mental.

In order for Gibb’s theory to work, a base neurological event corresponding to the targeted event must already be set to occur before the execution of any mental double preventer (think of the fair game scenario: the player must already be winding up to hit the bottle before the barrier is removed). But this seems counter to the notion of free agency, or even minimal mental causation. For example,<sup>85</sup> suppose that the neural processes have started

---

<sup>81</sup> Ibid.

<sup>82</sup> Ibid., 200.

<sup>83</sup> Ibid., 202.

<sup>84</sup> Ibid.

<sup>85</sup> What follows is adapted from Gibb’s own examples demonstrating her theory, Ibid., 203-206.

that will result in my messaging my leg to alleviate some muscle pain—these physical processes will be neurological event  $n1$ , my accompanying desire to alleviate my pain is mental event  $m2$ , labelled 2 because it is ontologically subsequent of the neurological event, thereby acting as a double preventer of any counterfactual manifestation that would put a stop to  $n1$ . But before I start body movement  $b2$ , the muscle pain goes away and I experience an intense itch in my arm that I want ( $m3$ ) to scratch ( $b3$ ). Unless there is a primed and ready neurological event,  $n0$ , that  $m3$  can act as a double preventer of  $n1$  for, I *must* rub my leg with the desire to scratch my arm. On the other hand, if  $n0$  is indeed present, it would have to be set in motion *prior to my forming a desire to rub my leg*, and even more troubling, *prior to my forming the desire to scratch my arm*.

Too much of Gibb's theory rests on the manifestation of powers permitting or excluding counterfactual events. Accepting this would create an ontological quagmire of objective, causally relevant, non-existent events—which may or may not be epistemologically accessible—that must be invoked when formulating a causal explanation that could attribute certain effects to particular agents. This seems *ad hoc* and, if taken seriously, impossible.

More recently, Matthew Owen has proposed mutual manifestation to specifically account for the physical manifestation of neurological activity that occurs with mental content. His goal is to refute neurological correlates as exclusive evidence for physicalism.<sup>86</sup> According to Owen, the physical body possesses passive powers that mental powers are active power partners with. For example, a runner chooses to run, and therefore she manifests a passive body power, or *B-Power*, along with an active mental power, i.e., an *M-Power*. These two powers, the *B-Power* and the *M-Power*, are in a relation of ontological

---

<sup>86</sup> Matthew Owen, "Aristotelian Causation and Neural Correlates of Consciousness," *Topoi* (2018): §1 and §3, <https://doi.org/10.1007/s11245-018-9606-9>.

dependence. Neither could manifest without the other—specifically, the manifestation of *B-Power* is a constituent of the manifestation of *M-Power*.<sup>87</sup>

I consider Owen's theory to be heading in the right direction when it comes to solving the mind-body problem for substance dualists. Of note is the constituent relation he proposes between mental and physical activities (i.e., that physical manifestations are a constitutive part of mental manifestations) since this potentially accounts for both physical causal closure and mind-body interaction. Given that physical manifestations are constitutive parts of mental manifestations, we have a built-in explanation for why physical manifestations accompany mental manifestations and, further, why mental manifestations have the ability to produce physical effects. However, Owen's view still leaves unexplained how a physically grounded power can have a mentally grounded power as a power partner for a single manifestation, regardless of whether that manifestation is consistent with the mutual manifestation thesis or not. This would be an important explanation to have to save substance dualism from the interaction problem, for the active power is just how a substance acts upon another substance with a passive power. Therefore, invoking a mentally grounded active and a physically grounded passive power still leaves a mental substance having causal power over a physical substance in need of explaining. That is, we are still left with a closure violation and the existing interaction problem, just at a deeper mereological level.

Moreover, Owen admits that the *B-Power* and *M-Power* are in a dependence relation. However, why should we consider the *M-Power* more ontologically primary in this dependence relation? Without a positive argument towards this end, we're left with an ontology compatible with physicalism, particularly one that is reflective of epiphenomenalism. We still have not been given reason to accept the ontological order of

---

<sup>87</sup> Ibid., §4.1.

dependency necessary for free will agency in a substance dualist ontology, other than that we should accept such an ordering to maintain free will (which is Owen's approach).<sup>88</sup> Even though I think Owen gets far with his theory, I believe there is still a long way to go.

#### **4.8 Conclusion**

What can we conclude after considering all these interaction theories? For one, the interaction problem is a difficult problem to get around: non-reductive physicalism, substance dualism, and, as shown above, even panpsychism seem on par in terms of faltering in the face of it. Idealism would be at an advantage at this point, except that it has its own problems in terms of accounting for truth and objectivity without having to invoke a divine observer.

Secondly, in light of the all the above theories (i.e., non-reductive physicalism, panpsychism, idealism, and substance dualism), it must be affirmed that those theories which allow for intrinsic properties are at a significant advantage. We know from chapter 1 that substance causation—involving real intrinsic causal powers—is more suitable than Humeanism at explaining what we causally observe. Moreover, from chapter 2, we have a reason to be committed to incompatibilist human freedom. Further, from chapter 3, we can make a case for the non-reducibility of experiential properties. Therefore, in order to preserve the idea that human beings have first-personal qualitative experiences—and, moreover, that they can rationally originate causal chains free from the determining effects of external forces—we have a strong reason to be committed to the existence of intrinsic properties.

Thus, physicalism must be rejected, including non-reductive physicalism, which does not allow for intrinsic properties, but merely breaks down to epiphenomenalism. This leaves us with the options of substance dualism, panpsychism, and idealism, all of which can allow for the existence of intrinsic properties in various forms. Some may have issue with the lack of importance I am placing on the existence of a distinctly physical realm. However, the

---

<sup>88</sup> Ibid.

existence of a distinctly physical realm is not integral to this thesis—only intrinsic properties are. Moreover, as Daniel Stoljar points out, defining *physical* as a distinct way of being is more difficult than one would think at first glance (see footnote 3 above).<sup>89</sup> We have ample reason to reject the modern conception of the physical, and, as Strawson argues, reason to admit what is historically considered nonphysical into the physical.<sup>90</sup> Instead of worrying about making difficult distinctions, our commitment should be to intrinsic properties, not to the existence of a physical realm interpreted in a narrow sense.

---

<sup>89</sup> Stoljar, *Physicalism*, 2-3, 28.

<sup>90</sup> Galen Strawson, “Realistic Monism: Why Physicalism Entails Panpsychism,” in *Real Materialism: and Other Essays* (Oxford: Oxford University Press, 2008), 53-54.

PART II:  
MORAL METAPHYSICS, KNOWLEDGE, AND  
PERCEPTION

## 5. The Question of Moral Perception

### 5.1 Introduction

What follows in the next three chapters is a defense of a strong moral realism that allows for non-inferential perceptual moral knowledge. In particular, the thesis I will be defending is that moral knowledge can be gained via ordinary perception (such as the knowledge of the presence of any ordinary object or properties, e.g., the green-leaved tree outside my window, can be gained via perception) and, moreover, that moral properties exist non-reductively and mind-independently as *relations*.<sup>1</sup>

The ideas developed in previous chapters are integral to the development of what follows. In chapter 1, I made the case that the relata of events are substances—as opposed to mere objects—and that these substances have real intrinsic features, including powers. This account of causation will be important to the account that I develop in chapter 6 of how moral relations are a type of causal relations between substances. It is important to this account that the substance theory of causation holds because, as I will argue in chapter 6, the associated powers and final causes of substances in a causal relation determine the moral quality—if any—of said relation. Without substance theory, there is no account of morality.

In chapter 3, I defended the idea that the experiential content of consciousness is phenomenal and non-reducible or eliminable (i.e., qualia realism is true). This will be important to the development of the idea that moral relations have their own phenomenology that is irreducible to descriptive properties or psychological states—a topic introduced in this chapter but dealt with more thoroughly in the next. Moreover, the experience of phenomenal moral relations is important to how moral concepts are learned and moral knowledge is gained, a topic of this chapter and chapter 7.

---

<sup>1</sup> More specifically, moral properties exist as *features of relations* between substances. What this means will be thoroughly explained in chapter 6. For the purposes of this present discussion, it is expedient just to say that moral properties are relations.

In chapters 2 and 4, I developed an account of the human being such that they are non-reducible to any material parts and possess intrinsic features and powers (i.e., exhibits incompatibilist causal freedom). Thus, I provided an account of the human substance which can enter into causal relations and, consequently, into moral relations. This will be important to the chapter 6 discussion of how and why human beings are morally culpable and are able to exercise moral agency, and thus why a substance ontology with human beings cannot help but be a moral one.

This chapter will serve as an introduction to my theory of moral perception as ordinary perception and moral properties as relational—as opposed to monadic—properties, with the focus of this chapter being on answering the question of whether or not moral knowledge can be gained via perception. In the next chapter, I will provide a metaphysical account of relations that is compatible with relations having their own irreducible phenomenology, and address the issue of how we can consider some relations moral. In the final chapter, I shall address the epistemological problems that arise from my view of perceptual knowledge, namely making sense of non-conceptual or concept-limited knowledge and how such knowledge is justified. It is the intention of these last three chapters to provide a viable metaphysical foundation for the moral realist while giving a plausible account of moral perception.

## **5.2 Harman's Cat**

Gilbert Harman, in the attempt to motivate a discussion on whether moral principles can be empirically analyzed, poses the following questions and thought experiment:

Can moral principles be tested in the same way [as scientific hypotheses], out in the world? You can observe someone do something, but can you ever perceive the rightness or wrongness of what he does? If you round a corner and see a group of young hoodlums pour gasoline on a cat and ignite it, you do not need to *conclude* that what they are doing is wrong; you do not need to figure anything out; you can *see* that it is wrong. But is your reaction due to the actual wrongness of what you see or is it

simply a reflection of your moral “sense,” a “sense” that you have acquired perhaps as a result of your moral upbringing?<sup>2</sup>

Ultimately Harman concludes that your reaction is from your moral “sense,” and not due what is objectively located in the world: people’s moral sensibilities determine what moral judgements they make, i.e., what they “see” in a given moral situation.<sup>3</sup> However, Harman does not arrive at the conclusion without first making some interesting observations about the nature of perception:

You see some children pour gasoline on a cat and ignite it. To really see that, you have to possess a great deal of knowledge, know about a considerable number of objects, know about people: that people pass through the life stages infant, baby, child, adolescent, adult. You must know what flesh and blood animals are, and in particular, cats. You must have some idea of life. You must know what gasoline is, what burning is, and much more. In one sense, what you “see” is a pattern of light on your retina, a shifting array of splotches, although even that is theory, and you could never adequately describe what you see in that sense. In another sense, you see what you do because of the theories you hold. Change those theories and you would see something else, given the same pattern of light... Observation depends on theory because perception involves forming a belief as a fairly direct result of observing something; you can form a belief only if you understand the relevant concepts and a concept is what it is by virtue of its role in some theory or system of belief.<sup>4</sup>

According to Harman, one’s theory of morality determines what moral judgements one makes in the same manner that one’s scientific theory determines how one may make sense of scientific observation.<sup>5</sup> However, unlike in the scientific case, moral judgements don’t necessarily involve assumptions about supposed moral facts in order to explain supposed moral observations, whereas with science, assumptions need to be made about physical facts in order to explain observations.<sup>6</sup> To support this claim, Harman argues that, since moral

---

<sup>2</sup> Gilbert Harman, *The Nature of Morality: An Introduction to Ethics* (New York: Oxford University Press, 1977), 4.

<sup>3</sup> *Ibid.*, 6-7.

<sup>4</sup> *Ibid.*, 4-5.

<sup>5</sup> *Ibid.*, 5-6

<sup>6</sup> *Ibid.*, 6.

facts do not play a role in how a subject arrives at a moral judgement (given that moral judgements have psychological explanations), then their existence can be doubted.<sup>7</sup>

Despite Harman's skepticism, there seems to be great intuitive appeal in the idea that a moral agent is just able to witness a moral situation unfold and "see" the morality of it, as Harman's cat thought experiment suggests—therefore, we have reason not to give up so easily on it. Moreover, Harman himself states that in order for us to see a cat being lit on fire by children, we need the concepts of *cat* and *children*, amongst other things. What if in order to see moral objects or properties, all we need to have is the correct moral concepts in the same fashion? Harman seems to argue that we do, but it seems that he would hold that these concepts do not denote anything beyond an agent's own psychology and thus do not project onto any moral facts in the external world.<sup>8</sup> The problem for the perceptual moral realist becomes locating the objects or properties that moral concepts latch onto—that is, moral referents—in objective reality. If we were able to accomplish this, it would provide us with an account of moral facts, and moral judgements would no longer be thought of as a mere product of an agent's psychology, but the product of perception.

### **5.3 Robert Audi's Theory of Moral Perception**

Robert Audi has done much to advance the claim that non-inferential moral knowledge can be gained through perception. Since he holds the widely accepted realist view that moral properties are constitutively grounded in more basic (e.g., physical or descriptive) properties, he argues that moral perception is the phenomenal experience of moral properties as perceived through the direct sensory phenomenal experience of ordinary properties

---

<sup>7</sup> Ibid., 7-8.

<sup>8</sup> Ibid., 4-7.

exhibiting certain patterns. Because it is indirect like this, moral perception isn't the same as ordinary perception for Audi, but something (imperfectly) analogous to it.<sup>9</sup>

To clarify his position, Audi makes an important distinction between a phenomenal perceptual kind of moral seeing from an intellectual kind of moral seeing. Intellectual seeing connotes more of an apprehension or understanding as opposed to just a mere perception.

Audi uses the example of seeing an American flag displayed every day in front of a residential home. Such a display could be “seen as” a political statement, and as such, *seen* in this instance conveys a type of apprehension or understanding—an intellectual kind of seeing. As an example of the perceptual kind of seeing, on the other hand, Audi offers the common illusion of a stick in water being “seen as” bent. In this instance, *seen* isn't conveying apprehension or understanding, but instead a purely perceptual phenomenon.<sup>10</sup>

With this distinction in mind, Audi makes clear he is interested in the perceptual kind of seeing and thinks moral knowledge can be gained through this kind of seeing.<sup>11</sup> In terms of Harman's concern over “seeing” moral properties and their grounding in reality, Audi's view allows for moral perception to be the perception of something that finds its ground in objective reality. His concern, along with my concern, is not merely with the internal psychological states of moral agents.

Audi further makes a distinction between what he calls “moral perception” and “perception of a moral phenomenon.”<sup>12</sup> The distinction is between an agent seeing a moral situation as moral versus her seeing the ordinary constitutive properties of a moral situation

---

<sup>9</sup> Robert Audi, *Moral Perception* (Princeton, NJ: Princeton University Press, 2013, 52-53, 57, 106-108; Robert Audi, “Moral Knowledge and Moral Perception,” *Proceedings of the Aristotelian Society, Supplementary Volumes* 84 (2010): 87-88, 94-95.

<sup>10</sup> Audi, *Moral Perception*, 9.

<sup>11</sup> Audi, “Moral Knowledge and Moral Perception,” 85, 87, 94.

<sup>12</sup> Audi, *Moral Perception*, 31.

but failing to see the situation as *moral*. For example, we could witness someone robbing a store, and all the nonmoral properties—e.g., prying open a door or shattering a window to get inside, the taking of mercantile goods, etc.—that such an action is constituted by. But even so, we could fail to see the *wrongness* of the act, i.e., the moral properties involved in robbery.<sup>13</sup>

This distinction between moral perception and perception of the moral finds its basis in Audi's metaphysical commitment to moral properties being grounded in base properties, with Audi calling the relationship that holds between moral and base properties both supervenient and consequential.<sup>14</sup> It is supervenient because "no two actions or persons can be alike in all their non-moral properties and differ in their moral ones,"<sup>15</sup> and consequential because the morality of a moral act is *determined* by descriptive or physical properties, i.e., the properties that make up the supervenient base.<sup>16</sup> Audi argues that supervenient or consequential moral properties themselves are not directly observable in the same way as base properties are directly observable.<sup>17</sup> Additionally, the phenomenal perception of moral properties does not map onto objects in the same way as ordinary properties seem to. In other words, moral perception is not pictorial or cartographical like visual perception. Audi argues this isn't *sui generis*—smell and taste do not seem cartographical either. And even though phenomenal sensory properties of experience may map onto the constitutive base properties of a moral action, they do not map onto moral properties themselves.<sup>18</sup> Nevertheless, Audi

---

<sup>13</sup> Ibid.

<sup>14</sup> Ibid., 35-36, 107-108.

<sup>15</sup> Ibid., 107-108.

<sup>16</sup> Ibid.

<sup>17</sup> Ibid., 35-36.

<sup>18</sup> Ibid., 37-38; Audi, "Moral Knowledge and Moral Perception," 88-89.

argues that moral perception involves “phenomenal representation constituted by a (richer) perceptual *response* to” moral properties than ordinary perception would have to ordinary properties.<sup>19</sup> It’s unclear what the phenomenal representation of moral perception itself exactly consists in, although Audi does mention that the constituents of moral perception may include psychological elements, such as attitudes and emotions, along with ordinary visual representations.<sup>20</sup> What is clear is that moral perception isn’t like ordinary perception, and, more important, moral perception operates, at least partially, through ordinary perception.

Importantly for our overall discussion, Audi affirms that moral concept possession plays an important role in what knowledge is gained via moral perception, and moreover, that moral perception aids in concept formation. The unexperienced subject can see the ordinary constitutive properties of a moral action without seeing the action as a moral action, and the reason that they may do this is because they lack the relative concept. However, acquaintance with the perception will help the subject in forming the appropriate concepts, and ultimately aid in forming moral knowledge. As Audi explains:

A child who has yet to develop the concept of injustice can see an act that constitutes an injustice. A decade later, of course, the same perception might immediately yield a moral conceptualization of the act or indeed moral knowledge thereof. Between these two points, the child may be disturbed at seeing an injustice in the kind of act in question, say giving a visibly much nicer pair of shoes to a twin sibling of the same sex taken to the same store just before the beginning of the school year. Prior to conceptualization, the child may have a sense of unfittingness in such action: the disparity in treatment disturbs the child who sees the sibling treated better. That perception of disparity, together with the sense of its unfittingness, reflects a discriminative sensitivity to differential treatment and puts the child in a good position to develop the concept of injustice. If this picture is correct, moral perception may precede, and indeed may be a normal developmental route to, moral concept-formation.<sup>21</sup>

---

<sup>19</sup> Audi, *Moral Perception*, 38; Audi, “Moral Knowledge and Moral Perception,” 88.

<sup>20</sup> Audi, *Moral Perception*, 48.

<sup>21</sup> *Ibid.*, 45-46.

This seems right and is along the same lines as Harman's own ideas on the subject, as discussed above. However, there is still the question of whether Audi has succeeded in giving an account of the properties moral concepts attach onto in the actual world. Here I think he has failed.

The issue concerns the constituents of moral perception. For Audi, these are not exclusively sensory phenomena, nor can they be. The reason being Audi's commitment to moral properties being supervenient and consequential—what is perceived by the subject is not moral properties *per se*, but their grounding. This supervenience theory is a very natural one for a moral realist to accept in the face of the extreme context sensitivity of moral properties. For example, a certain action (e.g., a *killing*) may in one context be morally impermissible (i.e., a *murder*) while in another be wholly permissible (e.g., in *self-defense*). Those theorists who hold to moral supervenience have a ready answer to why this is so: any change in grounding descriptive or physical properties leads to a change in moral properties. Moreover, it also seems obvious that moral properties cannot be the same as ordinary properties, like color or shape properties, since they cannot be “seen” in the same way as these properties. But it also seems obvious that moral properties must be possessed by objects (that is, *in* objects) in the same fashion as ordinary properties are, as Audi himself holds.<sup>22</sup> Yet, we don't *see* moral properties *in* objects. Therefore, the constituents of moral perception therefore cannot be purely sensory.

This leaves us with a theory of moral realism still liable to fail in light of Harman's objections. Given Audi's commitment to the supervenience theory of moral properties and perception (i.e., that moral properties themselves aren't sensed by the perceiving subject, but what is sensed is their nonmoral supervenient base) and subsequently his theory of perception

---

<sup>22</sup> Ibid., 48-49.

involving psychological states such as the “attitudinal, and even emotional aspects of moral perception,”<sup>23</sup> Audi’s theory is compatible with Harman’s own argument that moral judgments consist entirely in psychological states. There is nothing *per se* out in the world that moral concepts latch onto—the nonmoral base of moral properties may exist externally, but what one may consider to be supervenient moral properties may only be only a product of agential psychology, that is, his or her internal judgments. But what if we could meet Harman’s challenge by making the contents of moral perception purely sensory phenomena that directly reflect the way that external objects and/or properties exist mind-independently? Is this even possible? I argue that it is, if we reject supervenience theory as the basis for moral realism and instead view moral properties as relational properties. Before I get into the metaphysics of moral relations, however, I want to lay out a new theory of moral perception in light of this discussion of Audi’s theory.

#### **5.4 A New Theory of Moral Perception**

Along with Audi, I will also affirm the distinction between the *intellective* and *perceptual* senses of “seeing.” The notion that propositions of moral perceptual judgments can be thought of as an intellective kind of seeing is easy to accept and is difficult to deny. Consider an agent who says, “I see that this action is wrong.” In my view, we can understand this statement as akin to, “I see that two plus two equals four.” The agent is expressing a sort of intellectual understanding, that is, she is grasping that whatever action she is witness to is wrong. I do not deny that intellective grasping of moral truths is commonplace, or that perception can play a part in it. In accordance with Audi’s view, I will even admit it may be difficult to distinguish the two senses of seeing in some cases.<sup>24</sup> However, accounting for intellective moral grasping is not my concern here. My concern is whether moral knowledge

---

<sup>23</sup> Ibid., 48.

<sup>24</sup> Ibid., 9.

can be gained exactly like perceptual knowledge of ordinary objects. Therefore, my sole concern is with the perception in the sense of the pure qualitative perception of external mind-independent objects prior to any kinds of agential judgments. Hence, this is what I will be referring to in this discussion of perception and moral perception. Any kind of intellectual grasping is secondary to this kind of perception anyhow, as will become apparent throughout the discussion.

As I argued for in chapter 3, I also hold that perception—and now here, for the purposes of this discussion, moral perception—is phenomenal, i.e., it is subjectively experiential and there is a qualitative what-it-is-like to undergo it.<sup>25</sup> I also take perceptions of the same scene across different subjects—regardless of background knowledge, and barring any misapprehensions or sensory deficiencies—to have the same qualitative content. For example, one sees a red ball on a white table and might say, “I see a red ball on a white table,” or even, “I see *that* there is a red ball on a white table” (*that* being indicative of intellectual understanding).<sup>26</sup> However, neither being able to propositionally express one’s experience or have an intellectual grasp of undergoing said experience is necessary for awareness. An infant may see a red ball on a white table without having the concepts for *red*, *ball*, *white*, *table*, or *on* at all. Nevertheless, the infant would have the same visual experience—with the same qualitative content—an adult would have.<sup>27</sup> Moreover, both the infant and the adult could gain perceptual knowledge based on their perceptions.

---

<sup>25</sup> Audi would also agree with this. See Audi, *Moral Perception*, 12-13 and Audi, “Moral Knowledge and Moral Perception,” 81-82.

<sup>26</sup> See also Audi, *Moral Perception*, 10, for a discussion on how “seeing *that*” expresses one’s propositional knowledge.

<sup>27</sup> Although vision continues to develop throughout childhood and even adulthood, there is evidence that children’s vision is largely matured by age one. See Ruxandra Sireteanu, “Development of the Visual Field: Results from Human and Animal Studies,” in *Infant Vision*, ed. François Vital-Durand, Janette Atkinson, and Oliver J. Braddick (Oxford: Oxford University Press, 1996), 20; Anne Fulton, et al., “Development of Primate Rod Structure and Function,” in *Infant Vision*, ed. François Vital-Durand, Janette Atkinson, and Oliver J. Braddick (Oxford: Oxford University Press, 1996), 38-44; Russell J. Adams, “An Evaluation of Color Preference in Early Infancy,” *Infant Behavior and Development* 10 (1987): 149.

Many philosophers argue that propositional belief is necessary for any kind of knowledge and therefore an infant without the relevant concepts could gain no knowledge. This is especially true for those representationalists who take doxastic states to be contents of perception.<sup>28</sup> They would argue that an infant without the right concepts in fact fails to even see the objects said concepts are indicative of. Even Audi himself, who rejects the claim that representationalism entails this view about doxastic states,<sup>29</sup> would maintain that beliefs can only be held propositionally, including those beliefs that are held as perceptual knowledge.<sup>30</sup> Along with denying doxasticism (and representationalism in general), I deny that propositional beliefs are a necessary constituent of perceptual knowledge. The infant who sees a red ball doesn't need to hold the propositional belief *there is a red ball* in order to know there is a red ball. She has a kind of primitive non-propositional or extremely limited propositional knowledge—perhaps based on other imagery, primitive concepts, or gained wholly non-conceptually—of the object before her. She knows the ball is red, although she doesn't know it in terms of the *ball being red*, as an adult would. And even though the content of this knowledge isn't as detailed as an adult's, the infant can still act on it—that is, she can play with the ball.<sup>31</sup>

---

<sup>28</sup> Representationalism in general holds that perceptual content is comprised of intentional objects. By extension, representationalists tend to hold that what one believes about an object of experience determines the quality of said experience. See Chapter 3 section 3.3.1.3 for my previous discussion of representationalism.

<sup>29</sup> Audi, *Moral Perception*, 17-20.

<sup>30</sup> Robert Audi, *Epistemology: A Contemporary Introduction to the Theory of Knowledge*, 3rd Edition (New York: Routledge, 2011), 17-19. Audi argues that there is no particular proposition that needs to be believed by the subject in order for the subject to have knowledge of an object of perception. Therefore, the subject does not need to possess the correct corresponding concept of the object in order to have knowledge of said object. An example he gives is that of knowing that a field is *green* without knowing that it is particularly a *field*—perhaps one believes it is a rectangular plot of artificial turf. Thus, for Audi, propositions are always involved in beliefs.

<sup>31</sup> This will be more fully argued for in chapter 7. Audi has a similar view in that he distinguishes between *propositional* belief and *objectual* belief. The first is a belief in a particular proposition while the second is belief of or about an object of perception. The second requires no particular proposition to be held in order to be believed (see Audi, *Epistemology*, 18 and note 30 above). In both cases, however, some proposition needs to be held in order for belief to be possessed. I question this requirement.

Recall that Audi made a distinction between “moral perception” and “perception of a moral phenomenon,” i.e., between an agent seeing a moral situation as moral versus her seeing the ordinary constitutive properties of a moral situation but failing to see the situation as moral.<sup>32</sup> I make a similar distinction, but since I reject Audi’s metaphysics of supervenience,<sup>33</sup> the terms I use are *recognition* and *perception*, which better capture what I am distinguishing. In the instance where the agent sees the moral situation as moral, it can be said that she *recognizes* the situation for what it is—a situation exhibiting a moral quality. On the other hand, where the agent fails to see the situation as moral, it may be that she *perceives* the situation, even the moral quality of it, yet she fails to *recognize* the moral quality for the *kind* of thing that it is—a moral kind. Audi himself seems to also utilize the term “recognition” in the same manner; however, for Audi, recognition is dependent on the seeing of constitutive properties in a supervenience relation,<sup>34</sup> while for my view—since there is no supervenience relation—recognition just is knowing a property for what it is on sight. Thus, someone could fail to recognize, in my sense, a simple monadic property, such as the color red.

Thus, unlike the distinction Audi makes, my distinction between *recognition* and *perception* also applies in the nonmoral perceptual case.<sup>35</sup> The agent who perceives but fails to recognize a moral instance may do so in the same way an infant perceives a red ball but fails to recognize it as a *red ball*. The infant in this case fails because she lacks the concepts

---

<sup>32</sup> Audi, *Moral Perception*, 31.

<sup>33</sup> *Ibid.*, 107-108.

<sup>34</sup> *Ibid.*, 52-55.

<sup>35</sup> See Audi, “Moral Perception and Moral Knowledge,” 93, where Audi makes a distinction between *basic* and *non-basic* perception. According to Audi, basic perception is of basic properties, like colors and shapes, while non-basic perception is of things constituted by basic properties. One may object that my distinction between *recognition* and *perception* just is this distinction of his. However, I reject that basic perception, as Audi defines it, can be performed without what he would consider non-basic concept application, and so I reject his distinction.

*red* and *ball*. In the moral case, the agent may fail because she lacks the relevant moral concepts.<sup>36</sup> Note because the perceptual knowledge of concept-possessors is a result of recognition and not mere perception, the content of their knowledge will be more detailed and most likely propositional. For example, whereas the infant, with her limited or absent concepts, can gain perceptual knowledge of a certain colored object with a certain shape, the adult will have knowledge of a *red ball*, with all the accompanying knowledge of what balls are and what the color red is. Nevertheless, the quality of perception is the same for the adult and the infant. Also note that recognition requires (logically) prior perception, but perception does not necessarily entail subsequent recognition—thus, *recognition* can be thought of as a type of perception.

Some may object to the above by arguing that *recognition*—because it can result in complex propositional knowledge—is always a product of inference. Thus, there can be no such thing as non-inferential recognition, as I’ve defined recognition here. I reply that if we do not allow for non-inferential perceptual knowledge to obtain in complex—including moral—cases, then there is a question as to whether anything could be considered perceptual knowledge. This is because it can be argued that even the most basic qualitative properties, if propositionally believed, are inferred. For example, if I am to have a propositional belief that there is a *red ball* before me, then both the color *red* and shape *ball*<sup>37</sup> must be inferred on the basis of my qualitative experience. Thus, the condition that perceptual knowledge must be non-inferential if it is to be considered perceptual is impossible to fulfill. If we want to allow for non-inferential, perceptual knowledge in the basic case, there should be nothing barring

---

<sup>36</sup> Audi perhaps would agree with this assertion concerning concepts in the moral case, as mentioned above. However, it is important to keep in mind our moral metaphysics are drastically different. See Audi, “Moral Perception and Moral Knowledge,” 90.

<sup>37</sup> I am borrowing Audi’s notion of basic properties here. See Audi, “Moral Perception and Moral Knowledge,” 93.

the complex case. Audi himself argues similarly, claiming if our standard for perceptual knowledge includes only those things that cannot possibly be inferred, then, “[w]e could not be properly said to see that someone is angry or even that water is coming from a tap.”<sup>38</sup> Since we of course want to allow that we can know such things perceptually, we must reject the inferential standard that holds we can only perceptually know things that cannot possibly be inferred.<sup>39</sup>

There is also the subtly different concern of whether what I am calling *recognition* is really just an intellectual kind of “seeing”—i.e., merely a cognitive grasping—and not perceptual at all. This concern stems from the fact that *recognition* employs the use of concepts and, moreover, that concepts don’t seem perceptual in nature, but propositional and thus doxastic. Along with Audi, I do admit that the intellectual involves the phenomenal,<sup>40</sup> and therefore the demarcation between perceptual *recognition* and intellectual understanding is *prima facie* unclear (as already mentioned above). However, this demarcation problem between the purely perceptual and the purely propositional is not just a problem for me but for anyone who affirms the existence of perceptual knowledge and yet denies doxastic views of perception. If those who deny doxasticism want to allow for the existence of perceptual knowledge, then we cannot allow the deployment of propositionally expressible concepts during an instance of perception (thus, making the instance an instance of *recognition*) to bar us from claiming that a resulting belief is an instance of perceptual knowledge.<sup>41</sup>

Phenomenally and non-intellectively seeing a red ball as a *red ball* on sight should, at the

---

<sup>38</sup> Audi, *Moral Perception*, 52.

<sup>39</sup> See also Audi, “Moral Perception and Moral Knowledge,” 91-92.

<sup>40</sup> Audi, *Moral Perception*, 25; Audi, “Moral Perception and Moral Knowledge,” 85.

<sup>41</sup> Indeed, all theories of perceptual knowledge have this problem and the problem of inference above to contend with as possible objections, not just the one being developed here.

very least, be possible. Otherwise, we must either deny perceptual knowledge or affirm doxasticism.

As already mentioned, many philosophers assume that since we cannot *see* moral properties as we see ordinary properties, moral perception must not operate in the same manner as ordinary perception—Audi included.<sup>42</sup> I contend this view of moral perception is misguided on the basis that it assumes moral properties and their non-moral grounds are properties possessed by objects, a view that Audi himself is explicit about.<sup>43</sup> If moral properties were *in* objects like this, then necessarily they couldn't be the same as shape and color properties, since we don't see them in objects like these ordinary monadic properties. This view is reinforced given the context sensitivity of moral properties. A change in moral properties correlates with a change in physical or descriptive properties, and thus moral properties must, one might think, be supervenient. Given the different nature of moral properties, it is understandable that some might think moral properties cannot be perceived like ordinary properties.

In response, I maintain that moral properties are different from ordinary properties *not* because they are supervenient, but because they are not *monadic*: they are exclusively relational properties between substances. This relational account of moral properties also explains their context-sensitivity: a change in relata or in the nature of the relatum changes the moral quality of a moral relation (this idea will be further developed in the next chapter). Moreover, I contend these relations are sense-perceptible in the same manner as ordinary properties.

There is precedent for holding this view, namely from the world of music. Additionally, the examples that we can glean from music affirm that normative relational

---

<sup>42</sup> Audi, *Moral Perception*, 33-34; Audi, "Moral Perception and Moral Knowledge," 87-88.

<sup>43</sup> Audi, *Moral Perception*, 38-39, 48-49; Audi, "Moral Perception and Moral Knowledge," 88-90.

properties are sense-perceptible across all normal functioning subjects but take skill and experience to recognize (a notion that will become important over the next two chapters). Take tuning a violin for example: when a novice learns to play a violin, tuning may be a laborious exercise of playing each string individually, listening to each string's pitch and comparing it against the same note being played on a piano or some electronic device, and tightening or loosening the string many times over to match. However, when an expert violinist tunes a violin, all he needs to hear, if anything at all, is the pitch of an A4. Once he tunes his A string to this pitch, he is able to harmonize all the other strings around it, producing perfect fifths between them. He doesn't need to go up and down the scale to find a fifth up from A or a fifth below; he is able to hear the perfect fifth *itself* in the sounding of a chord between two adjacent strings because of the *quality of sound* that a perfect fifth is. This quality is dissimilar to two pitches played at random; it also seems to be something over and above than just the two individual pitches combined. It has its own what-it-is-likeness. Moreover, it is *recognizable* to the musical expert as a perfect fifth; indeed, even non-musical people can tell there is something qualitatively unique about it. What is being recognized as musically valuable or unique is a certain *fittingness* between two pitches—a relational property.

Some may balk at the above claim that there is something over and above the two pitches in a perfect fifth—how can I support this claim that there is this other *relational quality*? In response, consider the C major scale and E Phrygian mode: both C major and E Phrygian are comprised of A, B, C, D, E, F, and G natural. A melody in C major, however, will be qualitatively and drastically different than a melody in E Phrygian—yet, in each instance the individual pitches that constitute both melodies are exactly the same. The only

difference is how the pitches *relate* to each other in each melody. Changes in relation alter the perceptual quality.<sup>44</sup>

This, then, is my view of moral perception and moral perceptual knowledge, which will be defended over the next two chapters. During an instance of perceptual moral seeing, there is a perceptual *recognition* by the morally experienced of a relational property, due to the morally experienced being able to deploy the relevant moral concept or concepts involved on sight in the same manner she would deploy the concepts *red* and *ball* on sight if she saw a red ball. The reason the agent is able to *recognize* the moral relational property—as shared between a substance and itself, or between one or more substances—is because this relation has its own what-it-is-like quality to it. Like a perfect fifth, this quality derives from the relation being *fit* or *unfit* (as determined by the nature of the relata). Further, the recognition of this quality is sufficient for moral perceptual knowledge of what is perceived.

As such, when we are witness to a moral depravity, such as in the case of Harman’s cat, the recognition of the depravity—the kind of thing it is and why in this instance it is present—may come to us immediately, non-inferentially, and non-intellectively.<sup>45</sup> The less morally experienced may even be able to recognize that what she witnesses is *unfitting*, although she might not be able to put the quality into words or why there is a lack of fit. Audi himself even admits that the “unfittingness” of a moral wrong may be phenomenally “felt” by an agent. However, for Audi, this feeling is not always within the content of moral perception. For him, what essentially constitutes the content of moral perception is the non-

---

<sup>44</sup> This argument concerning relational phenomenology in both the visual and auditory case will be picked up again in chapter 6.

<sup>45</sup> That is, non-intellectively in the sense that a subject may not know why a situation has the moral quality that it does, or which features of the scene contribute to this moral quality. The subject will possess an understanding to the extent that she knows the scene before her is moral, so in that limited sense the recognition is intellectual.

moral base properties that ground moral properties.<sup>46</sup> On my view, however, the phenomenal unfittingness itself is, in this instance, the essential content of moral perception.

This is not to say perceptual recognition is the only avenue to moral knowledge. The perceptually deficient—i.e., the morally inexperienced—could infer her way to moral knowledge based on features of the moral circumstance. However, what is being maintained is that it is possible for moral knowledge to be gained in the same fashion as knowledge of a *red ball on a white table*—through ordinary perception. The way that this is possible is because 1) moral properties are relations, and 2) relations are perceptible. With this, we are able to map moral concepts onto real world entities, and thus meet Harman’s challenge.

---

<sup>46</sup> Audi, *Moral Perception*, 40.

## 6. The Metaphysics of Relations and Moral Relations

### 6.1 Introduction

In the last chapter I introduced a theory of moral perception that relied on the notion of moral properties as *relations*. In this chapter, I will modify this claim and develop a theory of moral properties as *features* of relations. Relations themselves, I argue, are the total complex state of affairs between substances, while their *features* are properties of that state of affairs. Moreover, relational propositions merely reflect *features* of relations as opposed to relations themselves. These distinctions are important, because—as it will be shown—it is only by making sense of relations in this fashion can relations be said to be sense-perceptible.

In order to clarify and defend this claim, I will provide a metaphysical account of relations in general and moral relations in particular. I will begin with a discussion of internal versus external relations. This is a natural starting point, for if both moral realism and substance realism are true, then *prima facie* the internal features of moral relata should strongly factor into moral relations (i.e., the natures of substances should determine what relations those substances stand in). As we shall see, determining whether or not relations are internal or external may help in determining how intrinsic properties of substances do factor into moral relations.

Moreover, if the phenomenology of moral features is above and beyond the monadic properties of relata (i.e., if moral properties have their own what-it-is-like quality) as I argue is the case, then at least some relations (particularly, moral relations) need to be demonstrated as non-reducible. Otherwise, moral relations cannot be considered sense-perceptible, and the theory being developed here lacks metaphysical support. This question is inextricably tied to that of the external or internality of relations, as shall be demonstrated below. I conclude that the assumption of non-reducibility being exclusively associated with the externality of relations creates a false dilemma: either relations relate intrinsic natures or they are non-

reducible, but never both. I argue that what philosophers typically consider to be internal relations are in fact non-reducible, and I thereby reject this dilemma. The reason that this can be so is because relations do not relate properties, but *substances*, which possess definite intrinsic natures. This allows for the non-reducibility and sense-perceptibility of moral relations.

Next I will move on to the specific question of moral relational features. What makes some features of a relation moral while others are not? I argue that moral features are a special kind of causal feature of relations. Moreover, I adopt the scholastic notion of final causes in order to account for the moral normativity of these special causal features. Thus, both the perceptibility and morality of moral properties are accounted for.

## 6.2 Internal Versus External Relations

According to David Yates, internal relations are “those [relations] whose holding is in some sense in the nature of their relata.”<sup>1</sup> The way most philosophers make sense of this concerns the location of truthmakers. If a relation is internal, then the truthmaker consists in the relevant monadic properties located in each of the relata.<sup>2</sup> In this way, the relation does not have a truthmaker external to the relata, and the relation itself can be reduced to monadic properties possessed by the relata—i.e., the relation is *reducible*.<sup>3</sup> For example, most philosophers hold that the monadic height properties of relata in a *taller than* relation is what makes the relation true, and thus the relation itself can be reduced to heights of the relata. An external relation’s truthmaker, on the other hand, is external to the relata, thus making the relation *non-reducible*.<sup>4</sup> Spatial relations—i.e., *next to*, *above*, etc.—are considered to be

---

<sup>1</sup> Anna Marmodoro and David Yates, “Introduction,” in *The Metaphysics of Relations*, ed. Anna Marmodoro and David Yates (Oxford: Oxford University Press, 2016), 7.

<sup>2</sup> *Ibid.*, 3, 7-8.

<sup>3</sup> *Ibid.*, 3.

<sup>4</sup> *Ibid.*

classic examples of external relations. Truthmakers for these relations are not found in monadic properties, but in relative spatial orientations of relata.

This way of thinking of internal versus external relations presents us with a dilemma. Given the commitment to a *substance* ontology with *intrinsic* properties being brought into this chapter, the notion that moral relations are internal relations seems intuitive—we want to say that the *natures* of substances, as expressed through intrinsic properties, factor into the moral relations they enter into. However, the reducibility of internal relations to monadic properties implies that moral relations, if they are internal, do not possess their own phenomenology—what we perceive in a moral instance is the monadic properties possessed by relata and nothing more. Therefore, if we are committed to the view that moral relations have their own phenomenology, we need to make a case for their non-reducibility and thus externality. However, if we do so, would not the natures of the substances involved in moral relations thus be left out of moral relations?

Perhaps the solution to resolving this dilemma lies in how we view the ontological relationship between what is generally considered to be internal relations and the associated monadic properties that they seem to arise from. Robert Knowles, in arguing for a view that allows physical objects to be related to platonic universals, suggests the following:

[A]ssociated with each relation are *relational properties* that hold of each relatum. For instance, the relation holding between *x* and *y* just in case *x* is taller than *y* has the associated relational properties *being taller than y*, which holds of *x* alone, and *being shorter than x*, which holds of *y* alone. The relation is not identical with these properties, but it is intimately related: the properties are instantiated just in case the relation obtains. Intuitively, the relational properties hold in virtue of the relation's obtaining, rather than the other way round. The relation is more fundamental.<sup>5</sup>

Even though I will not be defending Knowles' theory of universals as relata, I think there is something to be said for his ontological ordering of relations and relational properties. If the

---

<sup>5</sup> Robert Knowles, "Heavy Duty Platonism," *Erkenntnis* 80 (2015): 1258.

property of *being 6 feet tall* possessed by John gives John the additional property of *being taller than Jenn*, the property of *being taller than Jenn* is not prior to John's height relation to Jenn, but subsequent to it. The relation itself is more fundamental.

The fundamentality of a relation prior to relational predication seems to accord with Kit Fine's own analysis of the positional relations *on top of* and *beneath*:

Suppose that a given block *a* is on top of another block *b*. Then there is a certain state of affairs  $s_1$  that we may describe as the state of *a*'s being on top of *b*. There is also a certain state of affairs  $s_2$  that may be described as the state of *b*'s being beneath *a*. Yet surely the states  $s_1$  and  $s_2$  are the same. There is a single state of affairs *s* "out there" in reality, consisting of the blocks *a* and *b* having the relative positions that they do; and the different descriptions associated with  $s_1$  and  $s_2$  would merely appear to provide two different ways at getting at this single state of affairs.<sup>6</sup>

Thus, a single positional state of affairs exists between the blocks prior to predication, with any predication merely reflecting one of at least two relational descriptions.

Fine goes on to address the problem of how we can metaphysically and linguistically make sense of the single relation reflected by the single state of affairs.<sup>7</sup> Difficulty in doing this stems from the view that our relational predications, if true, must be made true by a truthmaker. If *a is on top of b* and *b is beneath a*, then it seems there must be two metaphysical entities fulfilling the truth conditions of these propositions. But how is this reconciled with only one relation?

In reply, Fine rejects what he calls "the standard view," i.e., the view that takes relations to hold on the basis of a certain ordering of relata.<sup>8</sup> The problem with this view is that it ontologically commits us to any relation under review (*a is on top of b*) and also its

---

<sup>6</sup> Kit Fine, "Neutral Relations," *The Philosophical Review* 109 (2000): 3.

<sup>7</sup> *Ibid.*, 4.

<sup>8</sup> *Ibid.*, 1-2. To clarify this ordering view, Fine provides the example of *loves*. If *a loves b*, then the relation *loves* holds between *a* and *b*, but not *b* and *a*, since *b* could fail to love *a*.

converse (*b is beneath a*).<sup>9</sup> Fraser MacBride explains how, with the standard view, the compounding of entities gets out of hand:

If we restrict ourselves to binary relations, converses might seem familiar enough because we have names for many of them and the grammatical device of passivization—‘after’ as well as ‘before’, ‘less’ as well as ‘greater’, ‘loves’ as well as ‘is loved by’. But we shouldn’t think that converse relations are familiar old friends. Each ternary non-symmetric relation has five mutual converses, and we don’t have names for any of them. Things only get worse when we consider that each quaternary non-symmetric relation has 23 converses, etc.<sup>10</sup>

Fine, however, rejects the idea that we have to accept the standard view. Because relations are mind-independent entities, he argues, it is difficult to make sense of them depending on any kind of ordering.<sup>11</sup> Instead, it seems more reasonable “to suppose that there is a single underlying relation connecting the things together and that any difference in the order of connection is to be attributed to the way we represent the relation as holding rather than to the relation itself.”<sup>12</sup>

Even though Fine understands the risk of multiplying entities unnecessarily, he still ends up with a theory of relations that does just that. It seems as if Fine thinks there is one underlying state of affairs that is exhaustively described by both *a is on top of b* and *b is beneath a*; any other relational description that can be applied to *a* and *b* requires another state of affairs.<sup>13</sup> Indeed, Fine is only concerned with addressing the problem of multiplying relations through their converses, not with how many relational propositions describe the same objects in relation.<sup>14</sup> This is why his theory is only *neutral* about the *ordering* of rela-

---

<sup>9</sup> Ibid., 2-3.

<sup>10</sup> Fraser MacBride, “How Involved Do You Want to Be in a Non-Symmetric Relationship?,” *Australasian Journal of Philosophy* 92 (2014): 4.

<sup>11</sup> Fine, “Neutral Relations,” 5.

<sup>12</sup> Ibid., 5-6.

<sup>13</sup> Ibid., 14.

<sup>14</sup> Ibid., 2-3.

he would consider *a is on top of b* and *b is beneath a* to be a single relation between *a* and *b* that is “neutral between *on top of* and *beneath*.”<sup>15</sup> But what if, in addition to the positional situation, *a* is dark in color and *b* is bright, let’s say, blue and yellow respectively? What underlies *a is darker than b* and *b is brighter than a*? A single state of affairs seems to make both *a is on top of b/b is beneath a* and *a is darker than b/b is brighter than a* true. Some may object that it is only the single *positional* state of affairs, which has nothing to do with color. But considering a relation like *the blue block a is on top of the yellow block b*, color and position seems descriptively inextricably linked. If we argue that they are distinct, then we lose the fact that, according to the state of affairs, both *darkness* and *blueness* is positioned on top of the *brightness* and *yellowness*.

Things only get more complicated if we add another relational feature, such as in *the small blue block is on top of the large yellow block*. In this case, there would be a multitude of relations: *block a is on top of block b*, *blue is on top of yellow*, *block a is smaller than block b*, *the dark block is on top of the bright block*, and etcetera. It seems clear that *the blue block is on top of the yellow block* and *the dark block is on top of the bright block* are predications that cannot reflect the existence of two different truthmakers. What I am further arguing is that *the small block is on top of the large block* and *the blue block is on top of the yellow block* cannot reflect the existence of two different truthmakers either. For indeed, if there were two different truthmakers, what then acts as the truthmaker for *there is more yellow than there is blue* given this state of affairs? Size then is also inextricably linked to color and position.

The problem with philosophers’ usual way of thinking about relations stems from thinking relations relate properties instead of substances. Indeed, a relation like *wider than*

---

<sup>15</sup> Ibid., 14.

seems to relate properties of width, while *taller than* seems to relate properties of height. But thinking of relations like this multiplies them needlessly. Instead, just as we think of width and height as properties of the same object, *wider than* and *taller than* should be considered properties—or *features*—of the same relation.

Some may object that this metaphysical picture of relations makes no sense of causal relations, an important metaphysical category of relations, given that causal relations are said to hold only when causal activity is occurring. This is especially pertinent for those philosophers who hold that causal relations hold due to the manifestation of corresponding powers intrinsic to individual substances (or objects, etc.) engaged in causal activity—it is *prima facie* unquestionable that the manifestation of causal powers is what does the relating when two substances are engaged in causal activity and not anything else. However, I answer that, when it comes to causal relations, any potential causal activity between two substances exists prior to the manifestation of any powers between them; the manifestation of powers is subsequent to, and a feature of, a prior relation that exists between two substances. In other words, what brings two substances into a causally relevant relationship is not the manifestation of powers *per se*: the manifestation of powers is evidence of a prior existent relation and is an instantiation of potential causal activity attributed to said relation. (Of course, this doesn't preclude that what determines the quantitative and qualitative causal features of a relation is determined by internal powers and capacities possessed by individual substances. However, these don't comprise the relation, since it is substances that enter into relations, not powers.) For example, a state of affairs that consists in a tablespoon of salt two feet away from a cup of unsaturated water possesses all the relational features relevant to causation: the salt is dissolvable by the water and the water is able to be made saline by the salt, even if no causal activity is happening. When activity is happening, this is merely a new

feature—one of many—brought about by a change in the state of affairs: perhaps now the salt has been dropped into the water.

Some may counter that I am overcomplicating the issue: isn't the manifestation and triggering of power between substances just what it means for them to be in a causal relation? In one sense, yes: when it comes to causation and causal explanation, the triggering of powers by other powers is what is relevant. Thus, when we typically talk of *causal* relations, we speak of two substances related by their causal activity. But in another sense, to require that the manifestation of powers be present in order for two substances to be considered in a causally salient relation would be to ignore a prior relation—with causally salient features—that exists between the substances already (e.g., the causally salient relationship that exists between salt and water even when water is not manifesting the dissolution of salt). There seems to exist a real causally relevant relation between substances that are not actively engaged in causal activity, a relation that can be expressed via predications concerning powers and liabilities possessed by said substances. As far as substances that are engaged in occurrent causal activity, the actual manifestation of powers is a product of this relation, not a precursor to it, ontologically speaking.

Now that we have a picture of what mind-independent relations really are, let us return to the discussion of internal versus external relations and the dilemma we faced. Recall that on one horn of the dilemma, we figured that our moral relations must involve the intrinsic properties—and thus the internal natures—of the substances involved in moral relations. Hence, we were compelled to think of moral relations as internal relations. On the other hand, we wanted to make the case that moral relations had their own phenomenology above and beyond the monadic properties involved in relations. Therefore, we were inclined to think that moral relations were non-reducible and, thus, external. The way in which this dilemma is resolved then is to disavow the false equivalence between a relation relating

natures and a relation being reducible to intrinsic properties: a relation can relate natures, not because it relates intrinsic properties, but because it relates *substances* that possess natures. Further, relations are not reducible because relations exist as the total complex state of affairs that substances stand in with respect to each other. This total complex state of affairs is not reducible to any one or more relata.

Consider for example the typical internal relation *taller than*. What grounds the *taller than* relation for most philosophers is the individual heights of the relata. Therefore, according to these philosophers, what makes the statement *Joe is taller than Mary* true is the fact that Joe is 72 inches tall and Mary is 62 inches tall. The relation is reducible to intrinsic properties and hence an internal relation. However, on the metaphysical picture being sketched here, this reduction would not be possible. Remember, it is not properties that enter into relations, but *substances*—a relation consists in the total complex state of affairs involving two (or more) substances (or between a substance and itself),<sup>16</sup> and relational predication merely distinguishes features of this relation. Thus, what makes *Joe is taller than Mary* true is the single total complex state of affairs that Joe and Mary find themselves in. This mind-independent relation exists in virtue of it being the case that Joe and Mary exist in the same universe—indeed, Joe would not be taller than Mary if Joe didn't exist. Therefore, what we consider to be the *taller than* relation is not reducible to the heights of either Joe or Mary; it is merely a feature of the non-reducible relation that exists between Joe and Mary.

It is not the case that monadic properties of relata have no role to play, however. If their heights were reversed, it would not be the case that Joe was taller than Mary. It must be the case that intrinsic properties determine the quantitative and qualitative features of a relation. Nevertheless, they do so not prior but subsequent to the existence of a relation.

---

<sup>16</sup> It is clear to me that a substance can be in relation with itself given a substance's ability to causally act upon itself. For example, I can willfully injure myself. This relation to oneself also accounts for why one is morally obligated to oneself, as we shall later see.

### 6.3 The Phenomenal Experience of Relations

Accepting the notion of non-reducible relations between substances still leaves us with the question of how such relations factor into our perceptions, especially since we don't seem to "see" these relations: what we seem to see is monadic properties. I question the denial that we do see these relations, however. If I am looking at two parallelly positioned inch-wide red rectangles drawn on a white sheet of paper, and I notice that one rectangle is longer than the other, that seems to me something that I can perceive directly. I would not be conjoining the perceptions I had of each rectangle's length; I could just see that one is longer than the other in a very brute and immediate fashion that suggests the irreducibility (to the phenomenal experience of monadic properties) of what I am seeing.

Arguably, the intuitive counter to the above assertion is simple: what I am seeing is the monadic phenomenal properties of each rectangle, especially since these properties factor so strongly into what is perceived. I do not deny that these properties do color the quality of my perception of the rectangles, so to speak. I just deny that those properties are *all* that I am seeing. What I am seeing is those properties contributing to the quality of a relation between the two rectangles, a quality that is distinct from the quality of each individual rectangle combined.

Some may object by pointing to an example of, say, two evergreen trees a thousand miles apart. Perhaps one is in California and the other is in British Columbia and, moreover, the one in California is 50 feet taller than the one in British Columbia. Can we indeed say that there's a real and irreducible relation between the two trees that has its own phenomenology, especially since we would never be able to feasibly see such a relation given the trees immense distance from one another? Even if we could fix cameras on both trees and compare the images of them side by side on a television screen, and if somehow this visual comparison could create an accurate visual representation of the size relation between the

two, this visual representation would only be momentary, and thus illusory—there is no real, enduring perceptible relation that exists between the two trees. At most, we should think of the relation that does exist as merely a logical or cognitive construct. So, since we can call into question the existence of a perceptible relation between these two trees, doesn't that call into question the existence of a perceptible relation between any two objects?

Whether or not a real, irreducible, perceptible relation exists between two objects at great distances seems to be the same kind of question as whether or not a tree that has fallen in the woods makes a sound. In both cases, there exists *something perceptible* that is nevertheless not perceived due to the limits of human perception. And, based on this fact, it is not reasonable to call into question the existence of the thing that would be perceived if circumstances, or human capacities, were different. I cannot hear a tree fall in my local wood because the wood is too far away for my human hearing. If I had the capacity for hearing at greater distances, such as an elephant does, I would hear the tree falling because there is indeed something to hear. In the case of the relation between the evergreens, if I had the capacity to somehow get two objects that are a thousand miles apart in my visual field at the same time, I would be able to see the relation that exists between them. There is also something to be perceived in that case as well.

Some may still be unsatisfied with the claim that a perceived relation is above and beyond the mere perception of each individual relata concurrently perceived. Perhaps another visual example can be used to demonstrate the notion that perceiving a relation is more than just the perception of each relatum's monadic phenomenal properties. Consider colored squares *A*, *B*, and *C*: in figure 6.1, we are able to see a relation of all three squares, while in figure 6.2 we can see a relation of just *A* and *C*.



Fig. 6.1

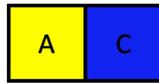


Fig. 6.2

The relation exhibited in figure 6.2 is able to manifest the contrasting nature of the colors comprising *A* and *C*—a contrast which produces a phenomenal experience that is exploited by artists—better than figure 6.1. This, even though in both figures, *A* and *C* are each plainly visible concurrently. The change in relation between *A* and *C* via either the absence or presence of *B* results in a perceptible qualitative difference between the two figures.

One may object that the qualitative difference comes merely via the presence or absence of *B*. Consider therefore colored squares *D*, *E*, *F*, *G*, *H*, *I*, and *J*:



Fig. 6.3



Fig. 6.4

In figure 6.3, the squares are arranged in chromatic order, while in 6.4, two squares—*E* and *I*—are switched. Yet, even though each and every colored square is concurrently perceived in both figure 6.3 and 6.4, and even though the change in relation is slight—only two squares switch places—the phenomenal difference between figure 6.3 and 6.4 is very apparent. The relation consisting of the colors in chromatic order is phenomenally above and beyond the colors merely being perceived concurrently.

An auditory example similar to the one presented in chapter 5 may well prove this point better than the above visual example. Consider the C major scale and the A minor scale.

Both C major and A minor are comprised of the pitches A, B, C, D, E, F, and G natural. However, a melody in C major will favor C as the tonic (i.e., “homebase,” or the pitch the melody finds its natural end on) while a melody in A minor will favor A as its tonic. A melody in C major will also sound more lighthearted while a melody in A minor will sound more serious. Why the difference between an C major melody and A minor melody? It is not merely because they are two different melodies—it is because of how the pitches relate to each other in each melody. The relation between pitches in one melody determines that C is favored as the melody’s end while the other determines that A is favored. The relation between pitches also determines that one melody has a lighthearted quality to it while the other melody has a more somber quality. Pitches play different roles—i.e., have different relations with each other—in a tune depending on whether it is in C major or A minor. For example, an E relates to an A in A minor as a dominant to a tonic, meaning that E has a naturally strong pull towards A. The relation between E and A in a C major melody, however, is of a mediant and submediant respectively. The pull going from E to A is not as strong in C major as it is in A minor. Changes in relation between pitches alter the perceptual quality of the tune above and beyond the mere phenomenal perception of the pitches that constitute the tune. To deny this is to deny large swathes of music theory and the common perceptual experience of listening subjects.

Some may have issue with the use of a musical example, considering the metaphysics of relations being developed in this chapter. Remember, according to the view being developed here, a relation is the total complex state of affairs that exists between two or more substances, or between a substance and itself. This state of affairs consists in a myriad of features, and relational predication (e.g., *taller than*, *on top of*, etc.) distinguishes individual features. So how do we metaphysically make sense of music, when music *prima facie* does not seem like a feature of a complex state of affairs between one or more substances?

It is important to recognize that, in the case of music, more than one feature of a relation is between presented to the listener. Not only pitch, but temporal distances and degrees of loudness or softness, and—in the case of live multi-instrument musical presentations—source of sound. This is indicative of the entities in relation being musical instruments (either taken to be objects in relation themselves or constitutional substances—such as wood or brass, etc.—in relation) or electronic pulses (such as those produced by a speaker). The relation between these entities has many features, not just ones having to do with sound. Indeed, a violin is *smaller than* a cello, and thus *smaller than* is another possible relational feature between musical instruments.

#### 6.4 The Metaphysics of Moral Relations

I now want to move onto the question of morality and relations. How can we classify some relations as moral? In other words, how can some *features* of a relation be considered *moral* features and how do we distinguish these features from non-moral ones? What properties of relata contribute to the moral *quality* of a relation? Answering these questions is critical to motivating the claim that moral properties are relational as opposed to monadic properties.

Consider the proposition *Joseph is taller than Jane*. Everyone would agree there is no moral truth (or falsity) being communicated here. Thus, Joseph being related to Jane by being taller than her does not seem on the face of it to be moral, and it would be difficult if not impossible to make the case that it is. On the other hand, consider the proposition *Joseph loves Jane*. Can we make the case that a moral truth could be communicated here? The answer is *prima facie* less obvious. Joseph's intentional attitude<sup>17</sup> could have moral consequences for his interpersonal relationship (or lack thereof) with Jane, but on its own,

---

<sup>17</sup> See E. J. Lowe, "There Are (Probably) No Relations," *The Metaphysics of Relations*, ed. Anna Marmodoro and David Yates (Oxford: Oxford University Press, 2016), 106-107, where Lowe makes sense of relational predicates like *loves* by appeal to monadic intentional states.

this proposition does not seem to qualify as morally evaluable. Thus, *Joseph loves Jane* does not reflect a moral feature of a relation.

Lastly, consider the proposition *Joseph murders Jane*: could this communicate a moral truth? Not *prima facie* in the same manner as the proposition *murder is wrong* could: *murder is wrong* blatantly expresses a moral judgment. But what about *Joseph murders Jane*? If indeed murder itself is morally neutral, then nothing moral is communicated with this proposition. But let us assume that murder is morally wrong, then *Joseph murders Jane* is morally communicative, and further, it seems to reflect a moral evaluation. Thus, *Joseph murders Jane* is reflective of a moral relational feature.

What features of the relation between Joseph and Jane does *Joseph murders Jane* capture that *Joseph loves Jane* and *Joseph is taller than Jane* fails to? I.e., what makes this feature of the relation moral as opposed to other features? On the face of it, *Joseph murders Jane* captures *causal* features. It expresses causal activity—the manifestation of powers—between Joseph and Jane. Joseph is exercising his powers to willfully and unlawfully cause the death of Jane, while Jane is manifesting potentialities to fatally respond to the powers Joseph is exercising. For indeed, if Jane was a granite statue, she would not possess said powers to be murdered. Thus, part of the quality of the relation *Joseph murders Jane* is dependent on the internal properties—the powers—of both Joseph and Jane.

It is apparent then that the moral features of relations are *causal* features of a special kind. Perhaps to see this more clearly, take for example again Harman's cat. In this example, we have a group of youths causally acting on a cat in a fashion so as to cause the cat pain and possible death. The cat is such a creature that it can be affected by the youths' actions. Involved in this moral situation are the causal agents of gasoline and a source of ignition, which the youths use to cause damage to the cat. Arguably, without causation, these entities

would lack the ability to enter into moral relations, and thus enter into any kind of meaningful moral relationships that would entail anything like obligations or possibility for harms.

The question then becomes, must causal activity always be present in order to say that a situation is moral? Indeed, we can think of situations where moral obligations may exist even when no causal activity is happening. Arguably, I have an obligation to be attentive to my infant child as she plays on the floor. I have an obligation to feed her when it comes time for her scheduled feeding. I have an obligation not to be abusive towards her. These obligations exist, if they exist, even though she's quietly playing on the floor and I watch her from the couch, and even though no causal activity is, at this instance, happening. If I fail at these obligations, by neglecting her or purposefully harming her, arguably this is a moral failure on my part. Therefore, the relation between my daughter and myself in this moment has moral features, even though no causal activity is occurring. But how can this be so?

Remember I argued above that causal features of relations exist, even though no causal activity is happening. Both my daughter and I possess powers and capacities to act on each other and to be acted on. These exist even when no power is being manifested. The existence of these powers, along with my knowledge of the harms and goods that can come about through their exercise, means that the relation I stand in with my daughter is a moral one. Therefore, moral obligations, if they occur, occur by default, arising merely by the existence of causally efficacious and intelligent creatures in a causal universe.

But do moral obligations occur? In the above example, it was assumed that I had a moral obligation to be attentive to my daughter, to feed her, and not to be abusive. But in order to motivate a robust theory of moral realism, an account must be given as to why I have these obligations. Or take for example again *Joseph murders Jane*. What makes this proposition indicative of a *wrong* moral relational feature? The *wrongness*—or *rightness*, for that matter—of a relational feature must somehow be metaphysically grounded. In other

words, accepting that moral features are causal in nature, there still is a question of what determines the moral quality of any causal relational feature and what distinguishes them from a non-moral causal feature. Otherwise, moral realism itself has not been demonstrated; just a causal realism that moral judgements can, real or unreal, be mapped onto. In my chapter 5 response to Audi's theory, I stated that, qualitatively speaking, the *fittingness* or *unfittingness* of a moral scene tells of its moral valuation. What, metaphysically speaking, can ground this fittingness or unfittingness? How can moral realism be grounded in a theory of moral relational features as causal features?

At the end of chapter 1, my chapter on general causation, I introduced the notion of final causes. The role that final causes plays in causation is that of "ordering" efficient causes, to borrow the wording of David Oderberg.<sup>18</sup> Final causes determine what causal effects are possible when a substance exercises its powers. Arguably, final causes play an essential and integral role in efficient causal activity,<sup>19</sup> determining the quality of the effects that take place.

I now want to make the case that final causes, particular to the natures of each and every substance, play an essential and integral role in determining the moral quality of moral relational features. For if moral features are causal in nature, then their quality—their fittingness or unfittingness—is determined by final causes. It is not my goal here to give a thoroughgoing account and defense of Scholastic metaphysics; my purpose here is merely to formulate a plausible account of how moral realism is grounded in causal activity between substances. Namely, I argue, along the lines of Oderberg, that (in scholastic terms) the actualization of potentialities towards perfection (being determined by a substance's nature)

---

<sup>18</sup> David S. Oderberg, "Finality Revived: Powers and Intentionality," *Synthese* 194 (2017): 2396; David S. Oderberg, *The Metaphysics of Good and Evil* (New York: Routledge, 2020), 28.

<sup>19</sup> Oderberg, "Finality Revived," 2396-2397.

is what constitutes a fitting moral action.<sup>20</sup> Unfitting action on the other hand hinders this process. In more analytic terms, an agent's actions are morally positive if the action in question is conducive to a substance or multiple substances trending<sup>21</sup> towards final states determined by their internal constitution, whereas wrong action has the opposing effect. This is the ground for moral realism I will defend. First, I will give an account of final causality. Afterwards, I will end the chapter by elaborating on what makes some causation moral.

#### 6.4.1 Final Causality

In Scholastic theory, a final cause is an end state towards which the exercise of power drives a substance, giving structure or order to causal activity, as well as determining its quality. Oderberg himself characterizes final causality as follows:

The Scholastics adhered to what is called the *principle of finality*: every agent acts for an end. The end, whatever it may be, is that which is perfective of the nature of the thing. Finality is another way of talking about *final causes*... Variations of the principle of finality include the proposition that every nature is ordered to an end; that nature does not act in vain; that the end is the first principle of activity; and that the end is the reason for all movement.<sup>22</sup>

What directs a substance towards this end is intrinsic to the substance itself, i.e., it is in the inherent *nature* of the substance—and the powers that it internally possesses—to naturally arrive at this end.<sup>23</sup> According to Oderberg, it is specifically higher-order causal properties that drive a substance to a teleological end, these higher-order properties being part of what

---

<sup>20</sup> David S. Oderberg, *The Metaphysics of Good and Evil* (New York: Routledge, 2020), 2.

<sup>21</sup> Oderberg uses the term *tending* to describe the same kind of causal movement by a substance, but I favor the term *trending* since it places emphasis on the final state as the goal of causal activity. It the movement towards or away from the final state that is important to the determination of whether a given situation is morally *fit* or *unfit*. See Oderberg, *The Metaphysics of Good and Evil*, 23.

<sup>22</sup> Oderberg, *The Metaphysics of Good and Evil*, 28.

<sup>23</sup> Edward Feser, *Scholastic Metaphysics: A Contemporary Introduction* (Piscataway, NJ: Transaction Books, 2014), 88-89. Feser also distinguishes here between intrinsic and extrinsic final causality in that an artifact will have its final causality imposed upon it from without, via an intentional agent giving purpose to the artifact, as opposed to from within an internal nature. Since I am only dealing in natural substances, which have only internal natures, I shall not be addressing extrinsic final causality.

Aristotelians and Scholastics call the *formal cause* of substance. Using the example of salt, Oderberg explains:

The reason for salt's dissolution in water is the final cause of its behaviour: salt is governed by a higher-order property in virtue of which it behaves in water in a certain way. That higher-order property is part of the *essence* of salt, what scholastics—following Aristotle's fourfold theory of causation—called the 'formal cause' of salt. In other words, it just is part of the essence of salt to be soluble in water: when we isolate any power or cluster of powers in virtue of which a substance behaves in a certain way, we are thereby isolating one or more final causes of the substance's behaviour.<sup>24</sup>

I will not utilize the Scholastic terminology of *formal cause*. Instead, I will borrow the notion of *final cause* capture the teleological nature of all causal activity along with the characteristic end states that substances seem to naturally arrive at through said activity. As Edward Feser points out, it is natural to an acorn to arrive at the end state of an oak tree: there is an internal, natural "directedness" of an acorn to arrive at this goal.<sup>25</sup> This is the reality I am wanting to capture via an appeal to *final causes*, since this reality will be important to the discussion of moral relational features as causal features. While in Scholastic theory (as Feser points out) the term *final cause* refers to only the external goal of this directness as opposed to the directedness itself,<sup>26</sup> for simplicity sake, I will use the term to capture all teleological aspects of causation that bear on substances, including their end states towards which their causal activity trends.

Many modern analytic philosophers, as well as scientists, are repelled by the notion of final causes. They assume teleology requires the existence of "spooky" properties that can only exist via supernatural means, thus running counter to any naturalistic explanation of universe. The main objectors along these lines would be those who hold to physicalism and

---

<sup>24</sup> Oderberg, "Finality Revived," 2397-2398.

<sup>25</sup> Feser, *Scholastic Metaphysics*, 88-89.

<sup>26</sup> *Ibid.*, 88.

the Humean view of causation, since they would reject the notion of intrinsic power properties altogether. I have already expressed a rejection of both reductive and non-reductive forms of physicalism, as well as Humean causation, in chapters 1 and 4. Along with not finding these theories plausible, I also take them to be incompatible with the moral realism being developed here. As such, I will not be addressing the physicalists' or Humeans' worries.

Panpsychists with commitments to naturalism, on the other hand, may also be hesitant to accept the strong teleological view being expressed here, but I don't think this needs to be much of a worry. Because panpsychism is compatible with intrinsic powers, there is nothing barring panpsychism from allowing this kind of teleology. As David Chalmers maintains:

The world...consists in a vast causal network of phenomenal properties underlying the physical laws that science postulates...The basic properties of the world are neither physical nor phenomenal, but the physical and the phenomenal are constructed out of them. From their intrinsic natures in combination, the phenomenal is constructed; and from their extrinsic relations, the physical is constructed.

On this view, the most basic laws will be those that connect the basic intrinsic properties.<sup>27</sup>

This view of Chalmers seems to be consistent with the idea that substances have internal powers that contribute to the directed nature of causation, however mentalistically the panpsychist wants to cash out these powers. Chalmers himself seems to want to cash out these internal properties in a strong mentalistic sense, as follows:

[I]t is often noted that physics characterizes its basic entities only *extrinsically*, in terms of their relations to other entities, which are themselves characterized extrinsically, and so on. The intrinsic nature of physical entities is left aside. Some argue that no such intrinsic properties exist, but then one is left with a world that is pure causal flux (a pure flow of information) with no properties for the causation to relate. If one allows that intrinsic properties exist, a natural speculation given the above is that the intrinsic properties of the physical—the properties that causation ultimately relates—are themselves phenomenal properties.<sup>28</sup>

---

<sup>27</sup> David J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (Oxford: Oxford University Press, 1996), 155.

<sup>28</sup> David J. Chalmers, "Facing Up to the Problem of Consciousness," *Journal of Consciousness Studies* 2 (1995): 217.

Nevertheless, even this strong phenomenal view of intrinsic properties is consistent with the existence of power properties if we take these powers to be products of intrinsic mental activity. Though this is a stark departure from most substance and powers theories, it is compatible with the notion of intrinsic properties of substances determining the quality of causal activity as well as the teleological nature of causal activity giving way to certain end states of substances. Therefore, panpsychism, even a panpsychism with a commitment to naturalism, is compatible with the notion of substantial natures and final causes—and thereby compatible with the moral ontology being developed here.

#### **6.4.2 Moral Causation**

The task now becomes one of making sense of the causal picture when it comes to the discussion of moral relational features as causal relational features. As already mentioned at the end of section 6.4, an agent's actions can be considered morally positive if the action in question is conducive to a substance or substances trending towards certain final states, the quality of which are determined by the internal constitution of substances, i.e., their natures. It is not my intention to demarcate the line between obligatory and supererogatory moral action, or even to identify any specific moral maxims. My sole concern is to develop a metaphysics in which the answer such questions can be ontologically rooted. As such, what follows, I believe, is compatible with multiple systems of ethics, although it seems better suited to some systems as opposed to others.

Also already mentioned is my borrowing from Oderberg's Scholastic theory, namely, his definition of *goodness as the actualization of potentialities towards perfection*,<sup>29</sup> which I am describing as the causal trending towards certain internally (i.e., via a substance's nature) prescribed (as opposed to causally determined) end states of substances. The natural causal

---

<sup>29</sup> Oderberg, *The Metaphysics of Good and Evil*, 2.

tendency towards *perfection*, i.e, the prescribed end state, is what constitutes *goodness*, i.e., a morally positive state of affairs. In Oderberg's words:

The older usage [of the term *perfection*]...sees perfection as what used to be called an *increase in the fullness of being*, a bringing to fulfilment or completion of some disposition, power, or tendency of an object—in Aristotelian-Scholastic terminology, the actualisation of some potentiality (or potency) of a thing. And this, at least according to the Scholastics, albeit less clearly in Aristotle himself, is precisely what they called *good* in the broadest, fundamental sense.<sup>30</sup>

Oderberg's definition of goodness—because it consists in the actualization of potency, and because for Oderberg all substances are a mixture of act and potency—even applies to non-organic materials.<sup>31</sup> However, I am uninterested in the case of non-organic materials, *per se*. Oderberg himself makes the distinction between a good state of affairs and a moral state of affairs, the latter requiring the presence and action (or lack thereof) of a moral agent.<sup>32</sup> What I am interested in is this latter case of a state of affairs involving a moral agent, i.e., a sentient being capable of moral knowledge and agency, although I am willing to allow that natural goodness that does not involve a moral agent, and perhaps goodness only involving inorganic entities. Because my main concern is with moral knowledge and agency, I will bypass the question of inorganic goodness and even organic goodness not involving a moral agent and focus on cases involving a moral agent. What is important for present purposes, however, is that the causal trending towards an end state—as determined by the nature of a substance—is what constitutes goodness.

When it comes to goodness which concerns a moral agent—specifically, in situations that involves other moral agents and other organic life, both sentient and non-sentient—the measure of what good action consists in can be described as action that fulfills the needs of

---

<sup>30</sup> Ibid., 15.

<sup>31</sup> Ibid., 57-58.

<sup>32</sup> Personal conversation with Oderberg.

the substances in relation with the moral agent (including the moral agent him or herself).

What these needs are depend on the essential natures of organic entities involved. According to Oderberg:

On the Scholastic view, a need is whatever is necessary for the proper functioning of some entity according to its nature. The nature is *presupposed*. Cows don't *need* to be mammals: anything that is not a mammal just is not a cow. Humans don't need to be rational, since anything non-rational is not human. But cows do need to lactate, to eat, to nurture their young; and humans need to *exercise* their rationality just as they need to exercise their bodies.<sup>33</sup>

Fulfilling these needs is what ensures causal trending towards the final states of substances.

Thus, good action consists in the fulfillment of these needs, and goodness itself consists in these needs being fulfilled.

Additionally, these morally pertinent needs arise via natures, which themselves are part of mind-independent reality. Oderberg makes this point as follows:

Now the kind of necessity appealed to here is broadly *metaphysical* in the following sense: all real needs must in some way emanate from an underlying proper function, that is, the functioning of an entity according to its nature. If there is no way of explaining how a putative need derives from an underlying proper function, it is not a need at all. It might be many things—a mere desire, a whim, a fancy, a requirement for this or that to happen—but it will not be a real need.<sup>34</sup>

For example, it is the proper functioning of a human child to grow into a healthy adult, and thus growing into a healthy adult fulfills the nature of a human child and can therefore be considered a good. This doesn't mean all needs are physical needs and that there are no felt needs or purely psychological needs. The example of human beings needing to exercise rationality, as Oderberg mentions, is a purely cognitive need. It also seems obvious that human beings have a need emanating from their nature for felt emotional connections to other human beings. I imagine there are other nonphysical needs that can be variously defined, and so we shouldn't limit what we consider to be real, metaphysical needs as those

---

<sup>33</sup> Oderberg, *The Metaphysics of Good and Evil*, 118.

<sup>34</sup> *Ibid.*

things that are merely necessary for basic physical growth and survival. But what should be clear is that goodness, as it is being defined here, is situated in metaphysically real, mind-independent reality.

If goodness comes by way of fulfillment, then moral badness—or evil—must come by way of privation,<sup>35</sup> or so it seems. A problem that immediately arises with this idea regards absences as causal entities since, when looking for a causal explanation for some evil, our first resort is to say that something necessary for a need fulfillment is lacking, such as a lack of water necessary for a plant to continue on living.<sup>36</sup> Some may be prepared to accept that absences have causal power. However, better metaphysical sense can be made of what seem to be causally potent “lacks.” Take for example respiration. We could say that a *lack* of oxygen causes death by suffocation, but what immediately kills a suffocating person is an accumulation of carbon dioxide waste in the blood. As for a *lack* of oxygen, oxygen is necessary for the renewal of metabolic energy and without it the body is subject to normal physical entropy. What causes death therefore isn’t a *lack* of anything: death is caused by waste accumulation and/or entropy, both of which are positive states of being (the waste build up itself having been caused by normal metabolic functioning).<sup>37</sup> Hence, if we define goodness as trending towards a designated final state, then we can define badness as trending away from that state. And even though we can make sense of goodness as fulfillment of natures, which causes final states to be trended towards, we don’t necessarily need to make

---

<sup>35</sup> Ibid., 3.

<sup>36</sup> Ibid., 175-176.

<sup>37</sup> See Ibid., 182-188. Oderberg would agree with this idea concerning real positive metaphysical causes in the face of seeming absences as causal entities (that is, he would deny that absences are real causal entities). Nevertheless, he seems to allow causal talk of absences—as causes *per accidens*—to be truth functional.

sense of badness as *lacks* in the metaphysically real sense, although it is explanatorily useful.<sup>38</sup>

Given the theory as now explained, we can make sense of the morality of a given situation as the causal relational features that either are conducive to or hindering to the trending of subjects towards their final causes. Let us consider an example, that of child drowning in a pool while in the presence of an adult who is capable of saving the child without any risk of harm to himself. For simplicity's sake, let's pretend these are the only two agents in this scenario, and no other relevant agents are present. Causally speaking, the water is interfering with the ability of the child's body to perform life sustaining functions (recall the analysis of suffocation above). Because there is such a risk of death, the child is not trending towards her final state—this comprises a badness or evil. The adult has—in accordance with his rational nature—knowledge, either through direct perception or practical reasoning, of the perilous state of the child. The adult also has—in accordance with his rational nature—knowledge of how to set the child trending towards her final state by removing her from the water. Additionally, the adult has the causal capabilities—and knows he has the capabilities—to set the child more trending towards her final state than she is at the moment: this comprises the causal relational features between the adult and the child.

Because the adult (1) has the causal power to save the child, (2) the knowledge of the causal circumstance the child finds herself in (i.e., her trending away from her final state), and (3) the knowledge of how he can cause her to trend towards her final state again, the adult is morally responsible in this situation. (1) is reflective of the causal relational features—moral relational features—of the relation that exists between the adult and the child. (2) and (3) are reflective of the nature of the adult as the kind of substance he is—a rational kind. Part of what makes him rational is his knowledge of the kind of substance the

---

<sup>38</sup> Ibid., 151.

child is, as reflected in (2). Thus, the causal relational features of (1) reflects a relation between substances as the kinds of things they are: a rational human adult with knowledge and capabilities and a human child that is, due to her nature, deserves saving in this situation. If the adult fails to act, he has committed a morally bad action. On the other hand, if he acts by taking the child out of the water, he has committed a morally good action. He is, because of (1)-(3), *morally* responsible and thus a *moral* agent.

A lot more can be said about the adult as a rational agent, such as what kind of knowledge we should really expect him to possess in the above scenario, and whether or not I am really holding him to too high a standard here (or perhaps too low) in terms of how much knowledge I think he possesses and should possess to be considered a moral agent. In other words, is it reasonable to expect a morally responsible agent to hold some particular knowledge given a certain scenario or should we say a rational agent is morally responsible regardless of what knowledge he or she possesses?

This is an interesting and important metaethical question, one that will not be fully fleshed out here. It seems apparent to me that a moral agent needs at least some working knowledge of the essential natures of the substances involved in a moral scenario, regarding how they causally interact. To put it differently, the moral agent needs knowledge of how his agency can affect change in his moral environment. Or, perhaps it is arguably the case that the moral agent *should* have such knowledge, and any lack of knowledge is a moral failure on his part. It may be the case that fully fleshing out an answer to this question will yield different answers according to which metaethical theory of agency is applied to the metaphysic described in this chapter. As such, this work will not be done here. The goal here is merely to lay out the basic necessary conditions for an agent to be considered a moral agent.

What is also important to emphasize is how these conditions are fulfilled by agents in the moral metaphysic developed in this chapter. Moral relational features—consisting of causal relational features of a special kind—are facets of relations that exist between substances. Moreover, what gives these relations the quality they have—moral or otherwise—is the monadic properties that exist in the relata. In the moral case, monadic essential properties comprise the natures of each of the relata, and thus determine the final causes that each relatum possesses. Trending towards, or away, from final causes determine whether or not a relatum is in a good state or a bad state. Further, how a moral agent relates himself—under his own causal control—to other relata determines whether or not he is committing a morally good action or a bad action.

## 7. Non-Conceptual Content and Perceptual Knowledge

### 7.1 Introduction

While discussing perceptual moral knowledge in chapter 5, I presented the case for a distinction to be made between *perception* and *recognition*. I had claimed that perceptual knowledge (of any kind, moral or otherwise) gained via mere perception was limited in scope, due to the lack of relevant concepts possessed by the perceiver, while perceptual knowledge gained via recognition was fuller, since it benefitted from the deployment of relevant concepts. In the cases of both perception and recognition, I argued that the qualitative content of perception was exactly the same. The only difference between the two cases was the deployment of relevant concepts *on sight* by the subject (or, in other words, the perceiver knows what the object or properties of perception are non-inferentially upon perception.)

In order to develop the distinction between perception and recognition, I had to make two *prima facie* problematic assumptions which I will now be dealing with in this chapter. The first is my assuming that the qualitative content of perception is wholly detached from concepts and can itself comprise knowledge. The idea that non-conceptual content exists is by itself a controversial position in perceptual theory. Can a plausible case be made for the qualitative contents of perception being experienced without the deployment of concepts? If the distinction between recognition and perception is to hold, the non-conceptual contents of perception must be coherently defined, and its connection to conceptual content explained. Thus, in this chapter, I will be defining non-conceptual content and bridging the gap between non-conceptual content and conceptual content. More controversial, however, is the idea that knowledge can be non-conceptual (i.e., non-propositional). (It is important to note that concepts and propositions are closely linked, for as Benjamin Bayer points out, “[o]ne does not use concepts except in the context of propositions, and one does not form a concept

except in the process of forming—and asserting—a proposition.”)<sup>1</sup> I will demonstrate the existence of non-conceptual knowledge by making a distinction between the knowledge possessed by pre-linguistic subjects—such as infants—and post-linguistic subjects, arguing that pre-linguistic subjects possess non-conceptual knowledge while post-linguistic subjects, for most things, do not (but are likely to for moral properties they do not possess concepts for). Key to this argument will be understanding how recognitional learning may occur given that perceptual content for pre-linguistic subjects is purely non-conceptual and thus non-discriminable via propositions. Features of perceptual content therefore must be discriminable perceptually—indicative of some kind of non-conceptual knowledge—if conceptual learning is to happen. This perceptual discrimination allows for concepts to be formed and attached onto discriminated properties. Otherwise, conceptual learning has no basis.

The second assumption is that a non-conceptual belief (a belief resulting from mere *perception*) is justified via immediate awareness of perceptual content alone, while a conceptual perceptual belief (one that involves some degree of *recognition*) is justified via a combination of immediate awareness and background knowledge concerning the deployed concept. How immediate awareness can confer justification demands explanation, especially given that the position that it can is also controversial in epistemology.<sup>2</sup> In order to tackle this problem, I begin with an exploration of Steven L. Reynolds’ work and entertain the notion that recognition, practiced as a skill, confers justification. Ultimately, however, after an exploration of the work of Carl Ginet and William Alston, I argue that immediate awareness of the content of perception itself directly confers justification when it comes to mere

---

<sup>1</sup> Benjamin Bayer, “A Role for Abstractionism in a Direct Realism Foundationalism,” *Synthese* 180 (2011): 370.

<sup>2</sup> Laurence Bonjour, “The Regress Problem and Foundationalism,” in *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa (Malden, MA: Blackwell Publishing, 2003), 17-19.

*perception*, while the combination of immediate awareness and background conceptual knowledge confers justification when it comes to *recognition*.<sup>3</sup>

Dealing with these assumptions will demonstrate that in both the cases of perception and recognition, knowledge is justified non-inferentially, and even though this seems like a jarring departure from the moral discussion we been having up to this point, dealing with the justification issue is integral to my theory of moral perceptual knowledge. This is because an intuitive objection to my view that moral relational features have their own irreducible phenomenology (i.e., a phenomenology over and above the monadic properties possessed by substances involved in a moral situation) would be that, because moral subjects don't always come to the same moral evaluations when presented with a moral scene, irreducible moral phenomenology must not exist; therefore, moral evaluations in response to perceptual scenes must always come by way of inference when they do come. My answer to this objection is that, like pre-linguistic subjects, those who lack the relevant moral concept (or concepts) for a moral scene can still be said to perceive the moral phenomenology of the scene but fail to recognize it (thereby accounting for some vary in evaluations). On the other hand, those who possess the relevant moral concept and are able to deploy it on sight (i.e., those who are able to *recognize* the moral scene for what it is) are justified in their moral evaluations. This chapter explains how perceivers but especially recognizers can be justified in their perceptual moral beliefs without the need to be justified via inference, thereby allowing me to defend my view against this objection.

Of importance to keep in mind in this chapter is a commitment to a robust view of qualia realism, which was defended against eliminativist and representationalist views of phenomenal experience in Chapter 3. I will here touch on representationalist arguments once

---

<sup>3</sup> An internalist view of justification is being assumed throughout. I will not be defending internalism against externalism.

again, but this time in the context of an epistemological discussion concerning the content of perceptual awareness and perceptual knowledge, and thus much will be assumed against representationalism. For arguments concerning representationalism as a theory of phenomenal experience, see Chapter 3, section 3.3.1.3.

## 7.2 Conceptual and Non-Conceptual Content

Concerning the relationship between concepts and experience, Matthew Burstein asks, “Do our experiences yield direct knowledge of the world, or do they do so only when the experiences are mediated by concepts?”<sup>4</sup> This question concerns the contents of perception—whether it is conceptual, non-conceptual, or a mixture of both. Answering this question is of particular epistemological importance since the answer can shed light on how our perceptions and perceptual knowledge relate, especially in regard to the question of how one—our perceptions—confers justification to other—our perceptual beliefs.<sup>5</sup> The question of justification will be dealt with in section 7.3. My first concern is with defining non-conceptual content and how it relates to conceptual content, demonstrating that non-conceptual content exists as distinct from conceptual content, and making the case for non-conceptual knowledge.

Non-conceptual awareness can be thought of as the *unmediated* (i.e., immediate) awareness of mind-independent entities, i.e., an awareness of external objects that is *given* directly to one’s conscious awareness.<sup>6</sup> In other words, non-conceptual content, if real, would lack mediation through concepts or propositions, and thus would involve a conscious awareness consisting purely of the qualities that constitute the perception of an external

---

<sup>4</sup> Matthew Burstein, “Situating Experience: Agency, Perception, and the Given,” *The Southern Journal of Philosophy* 44 (2006): 2.

<sup>5</sup> Ibid.

<sup>6</sup> William P. Alston, “Back to the Theory of Appearing,” *Philosophical Perspectives* 13 (1999): 182.

object.<sup>7</sup> However, considering the question of how non-conceptual content could hold in the subject's mind (i.e., how the subject could be consciously aware of the content), there are many who deny its existence and/or its significance. Tim Crane, who maintains that concepts are required for intentional states, argues that there is no such thing as thoughts without concepts; thus, without possessing concepts, a subject couldn't grasp anything.<sup>8</sup> Fodor, on his part, argues that non-conceptual content, if it existed, couldn't carry any ontological commitments.<sup>9</sup> This is at least partially due to the fact that non-conceptual representation lacks the power to individuate objects due to the absence of quantifiers.<sup>10</sup> Ernest Sosa goes as far as to deny that there can be subjective awareness of properties without concepts; for him, there is a tight connection between perception and concept possession.<sup>11</sup>

Moreover, it is a widely held view that, even if non-conceptual awareness exists, it cannot, without attaching to a proposition, comprise a belief. In other words, there is no such thing as a non-conceptual belief and, consequently, no such thing as non-conceptual knowledge. Even Alston, himself a major advocate for the non-conceptual awareness of the subject, makes the explicit claim that non-conceptual awareness does not necessarily result in the awareness of any facts.<sup>12</sup> Thus, he detaches the notions of non-conceptual content and belief, and affirms there is no such thing as non-conceptual knowledge.<sup>13</sup> Along the same

---

<sup>7</sup> William P. Alston, "Sellars and the 'Myth of the Given,'" *Philosophy and Phenomenological Research* 65 (2002): 71-72.

<sup>8</sup> Tim Crane, "The Nonconceptual Content of Experience," in *The Contents of Experience: Essays on Perception*, ed. Tim Crane (Cambridge: Cambridge University Press, 1992), 144-148.

<sup>9</sup> Jerry Fodor, "The Revenge of the Given," in *Contemporary Debates in Philosophy of Mind*, ed. Brian P. McLaughlin and Johnathan Cohen (Malden, MA: Blackwell Publishing Ltd, 2007), 109

<sup>10</sup> *Ibid.*, 109-111.

<sup>11</sup> Ernest Sosa, "Does Knowledge Have Foundations?" in *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa (Malden, MA: Blackwell Publishing, 2003), 124-125.

<sup>12</sup> Alston, "Back to the Theory of Appearing," 183.

<sup>13</sup> Alston, "Sellars and the 'Myth of the Given,'" 70.

vein, Audi argues that, while perception doesn't depend on conceptual content, belief does. The example he gives is that of a rectangular field: a subject that lacks the concept *field* would perceive a green field and its shape, but would not be able to form a belief about a *field*.<sup>14</sup>

This stance is due to an assumption that non-conceptual content is not the kind of thing that can be true or false.<sup>15</sup> Additionally, it is assumed that it is propositional content, not non-conceptual content, that requires justification.<sup>16</sup> Without this requirement to be justified, non-conceptual content cannot obtain the status of knowledge, i.e., true *justified* belief. As John Beversluis argues in critique of John Baillie, committing oneself to non-propositional knowledge commits one “to the awkward position of maintaining that one’s knowledge *of* any object is not only logically, *but also epistemologically*, prior to any affirmations” one could make.<sup>17</sup> Again, the assumption is being made that affirmations, in the form of propositions, are necessary for knowledge.

Setting aside for a moment the question of non-conceptual knowledge, can a case be made for the existence of non-conceptual content? It can be when considering how it is that concepts are applied. As Alston contends, the non-conceptual content of perception is fundamental to perception in that “the deployment of concepts is based on it and presupposes it.”<sup>18</sup> Likewise, Bayer maintains that the non-conceptual (the word he uses is “preconceptual”) ability to see resemblance, i.e., that two objects resemble each other in

---

<sup>14</sup> Robert Audi, *Epistemology: A Contemporary Introduction to the Theory of Knowledge*, 3<sup>rd</sup> Edition (New York: Routledge, 2011), 19.

<sup>15</sup> Bayer, “A Role for Abstractionism in a Direct Realism Foundationalism,” 358-359.

<sup>16</sup> *Ibid.*, 359.

<sup>17</sup> John Beversluis, “Some Remarks Concerning a Non-propositional Knowledge of God,” *Australasian Journal of Philosophy* 43 (1965): 378

<sup>18</sup> Alston, “Back to the Theory of Appearing,” 186.

some capacity, is required in order to form a concept.<sup>19</sup> The necessity of non-conceptual content in learning concepts seems right, for indeed how would concepts latch onto the external world if there was nothing perceived to be latched onto? In the instance of a child's perception, there would be nothing about occurrent experience onto which to place a concept she is in the process of learning—she would be unable to recognize the concept visually presented or represented in the content of her perception.<sup>20</sup>

As Alston points out, a commitment to the notion that perceptual experience is essentially non-conceptual is not necessary in order to admit non-conceptual content.<sup>21</sup> In fact, Alston holds the view that adult perception relies heavily on the deployment of concepts.<sup>22</sup> The notion that adult perception is largely conceptual seems correct considering that, given adult background knowledge, adult perception would be largely *recognition* in comparison to a child's.<sup>23</sup> Even so, a case can be made for the existence of non-conceptual content even in adult *recognition*. Non-conceptual content is what would make adult perception of external objects markedly different from mere thoughts about external objects.<sup>24</sup> As Alston points out, visually experiencing a tree drastically differs qualitatively from merely possessing or forming a belief about a tree.<sup>25</sup> Likewise, thinking about all the entailing propositions concerning an object differs from visually experiencing said object.<sup>26</sup>

---

<sup>19</sup> Bayer, "A Role for Abstractionism in a Direct Realism Foundationalism," 369.

<sup>20</sup> Crane, "The Nonconceptual Content of Experience," 137.

<sup>21</sup> William P. Alston, "Perceptual Knowledge," in *The Blackwell Guide to Epistemology*, ed. John Greco and Ernest Sosa (Malden, MA: Blackwell Publishers, 1999), 232.

<sup>22</sup> Alston, "Sellars and the 'Myth of the Given,'" 73.

<sup>23</sup> See also Alston, "Sellars and the 'Myth of the Given,'" 73.

<sup>24</sup> Alston, "Sellars and the 'Myth of the Given,'" 72-73. See also Laurence Bonjour, "Back to Foundationalism," in *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa (Malden, MA: Blackwell Publishing, 2003), 71-72.

<sup>25</sup> Alston, "Back to the Theory of Appearing," 194.

<sup>26</sup> Alston, "Sellars and the 'Myth of the Given,'" 72.

The difference between visually experiencing an object versus merely thinking about it propositionally is due to the non-conceptual content involved in the experience, which gives it a distinctive quality.<sup>27</sup>

Furthermore, as Alston argues, the fact that the distinctive quality of an experience can vary in specificity in comparison to propositional content is due to non-conceptual content and propositional content being distinct.<sup>28</sup> The example Alston uses is shades of red. An object can look red without looking any kind of specific shade to the observer. Although every case of looking red involves some specific shade, the observer may fail to notice the specific shade or lack the appropriate concept.<sup>29</sup> Bonjour likewise points out that the variances in shading, lighting, and coloring—although perceptually present to the conscious awareness of an observer—cannot be conceptualized, and therefore what is perceived can be more refined than what can be expressed.<sup>30</sup> The opposite can also be true:<sup>31</sup> we can come up with an infinite amount of propositions about a visual scene, given all the objects, properties, and relations involved. However, the non-conceptual qualitative content isn't infinite, but limited.

Even given the existence of non-conceptual content, can a case be made for non-conceptual belief? It can when we consider subjects' varying linguistic skills and the content of intentional states. While arguing for the position that concepts are necessary for belief, Crane utilizes the example of a subject who believes she is 70 years old. In order for the

---

<sup>27</sup> Ibid., 73.

<sup>28</sup> Ibid., 77.

<sup>29</sup> Ibid., 77-78.

<sup>30</sup> Bonjour, "Back to Foundationalism," 71-72.

<sup>31</sup> Alston, "Sellars and the 'Myth of the Given,'" 77.

subject to possess this belief, Crane maintains, she must also possess the concept *year*.<sup>32</sup> However, maintaining this tight relationship between concept and belief isn't as straightforward as this case seems to suggest. Take for example a small child's concept of a *year* versus the same concept possessed by an adult. It is commonly known that children experience time more slowly than adults (and young adults more slowly than older adults) and thus a child's concept of *year* would be that of an exceedingly long time. Moreover, a small child's concept of a *year* would lack robust background knowledge and fully fleshed out ideas related to how a *year* is calculated and societally experienced (e.g., the earth revolving around the sun, seasonal changes, cultural work and holiday rhythms, fiscal calendars, the collective cultural experience of the current zeitgeist in comparison past histories, etc.). Thus, a child's belief that grandma's 70 years old versus an adult's belief that grandma's 70 years old may have wildly differing content. Still, this fails to demonstrate that beliefs can be *conceptless*. In order to demonstrate this, we must look at what role concepts and propositions play in belief formation.

Alex Grzankowski, who is strongly against the non-conceptual view of perceptual content, defines propositions as "abstract truth-evaluable entities that are the objects of the propositional attitudes such as belief and desire and that are expressed by declarative sentences."<sup>33</sup> With this in mind, Grzankowski attempts to argue that representational content—such as the content of perception given representationalism—must be propositional (and thus, conceptual) since the content is linguistically expressible.<sup>34</sup> However, the question remains whether the contents of perception *must* be expressible by the subject that is

---

<sup>32</sup> Crane, "The Nonconceptual Content of Experience," 141.

<sup>33</sup> Alex Grzankowski, "Pictures Have Propositional Content," *Review of Philosophy and Psychology* 6 (2015): 152.

<sup>34</sup> *Ibid.*, 151-152.

undergoing a perceptual experience, i.e., whether the attitude of said subject must be *propositional*, or if it is enough that the attitude *would* be if it was possessed by a subject that had the appropriate linguistic skills.

Even so, Grzankowski goes on to argue that those subjects who lack certain language skills nevertheless possess the ability to express *some* content, and therefore it cannot be said perceptual content is *not* propositional. The example he uses is of Piraha speakers who possess limited numerical terms and therefore cannot express that there are *fifteen* coins lying on a table. Grzankowski concludes that, even though there is this expressive deficit, this doesn't entail that Piraha speakers can't propositionally express *some* things about the coins on the table, and we can therefore conclude that perception is wholly propositional.<sup>35</sup> Grzankowski draws the wrong conclusion from this example, however. It is plain that Piraha speakers see the fifteen coins on the table just the same as English speakers would—what would qualitatively be presented or represented to them does not differ. Nevertheless, since the Piraha speakers lack the appropriate numerical concepts, they are unable to *express* what they perceive in the same fashion as English speakers. If that same perception was had by an English speaker, it *would* be propositionally expressible by the subject whose perception it was. The Piraha's perceptual belief about the coins does not merely consist in the limited propositional content that they can express, as Grzankowski would argue. It also consists in the phenomenal content that they *cannot* express, that would be expressible *by them* if they possessed the appropriate concept of *fifteen*.<sup>36</sup>

---

<sup>35</sup> *Ibid.*, 157.

<sup>36</sup> Carl Ginet argues that beliefs that are inexpressible by a subject who lacks linguistic ability are nonetheless *propositional*. This position, however, seems to muddle the issue since it doesn't account for the difference in content between an expressible and inexpressible belief. See Carl Ginet, *Knowledge, Perception, and Memory* (Dordrecht, Holland: Reidel Publishing Company, 1975), 5-6.

This seems to put into question the role that concepts or propositions play in beliefs. What *is* the role of concepts in forming and maintaining beliefs? Crane holds that concepts are constituents of intentional states, intentional states being the objects of inference.<sup>37</sup> If this is the role concepts play, we can ask if something else can play this role. Indeed, we can ask whether intentional states must be propositional, and whether the role that concepts play in intentional states can be played by a non-propositional entity other than concepts.

Arguably, there is such an entity: the unmediated experience of objects themselves.<sup>38</sup> My experience of seeing a book on the coffee table is about the book on the coffee table, given that I give appropriate focus to the book on the coffee table, regardless whether I know what books or coffee tables are. And even though this intentional state lacks the concept *book*, I am not arguing that it results in the sophisticated and expressible knowledge that we find in propositional knowledge. At the very least, it is hard to deny that unmediated experiences would result in at least one belief: the belief that an object *is*, i.e., that it is in the presence of the subject of experience. If such a conceptless belief exists, perhaps there can be others.

Even if we want to deny that purely conceptless beliefs exist, it is difficult to deny that possession of concepts involved in propositional belief, and the subsequent knowledge, comes in degrees. Utilizing an example from Audi, a little girl may not know what the object on the table is (a *tachistoscope*), but she knows that whatever *it* is, *it is noisy*.<sup>39</sup> This is not indicative of the child failing to see the object on the table and thus failing to gain knowledge of it. What she fails at is deploying the relevant concepts, which she lacks. Furthermore, even

---

<sup>37</sup> Crane, "The Nonconceptual Content of Experience," 147.

<sup>38</sup> See also Alston, "Sellars and the 'Myth of the Given,'" 72, where Alston admits that direct awareness is intentional in the intuitive sense of being about something, but seems to deny—although it is left unexplained why—that it can fully be considered intentional.

<sup>39</sup> Audi, *Epistemology*, 19-20.

though her non-conceptual knowledge may be crude, it isn't qualitatively limited. She has the same qualitative experience as a concept-possessor. Otherwise, she would fail to apply the appropriate concept to the object on the table once she possesses it.

Regardless of the above arguments, it is impossible to deny outright that pre-linguistic children and some animals gain perceptual knowledge of their surroundings.<sup>40</sup> This is because their behavior tells us as much.<sup>41</sup> Ginet argues that beliefs are dispositional states, and as such, their presence can be tested by observing a subject's physical or cognitive behavior.<sup>42</sup> Even though we may not want to so bluntly equate beliefs with dispositional states as Ginet does, we can acknowledge a dispositional component to some beliefs that can manifest itself in behavior. Since we know that animals and infants can successfully navigate their environment,<sup>43</sup> we can conclude animals and infants possess knowledge.

Given this, then perceptual content and perceptual knowledge could consist in conceptless content exclusively in the pre-linguistic.<sup>44</sup> We can accept this while also accepting that knowledge in the post-linguistic is largely conceptual. It is not unprecedented to think that perceptual knowledge differs between subjects in this fashion. John McDowell argues that the perceptual knowledge of rational subjects and the perceptual knowledge of non-rational and pre-linguistic subjects (i.e., animals and small children), although fundamentally the same, differ in that the warrant for belief possessed by each is distinctive.<sup>45</sup>

---

<sup>40</sup> William P. Alston, "What's Wrong with Immediate Knowledge?" *Synthese* 55 (1986): 90.

<sup>41</sup> Ginet, *Knowledge, Perception, and Memory*, 25. See also William P. Alston, "Level Confusions in Epistemology," in *Epistemic Justification: Essays in the Theory of Knowledge* (Ithaca, NY: Cornell University Press, 1989), 164. Alston takes this kind of knowledge to be mediated whereas I take it to be immediate.

<sup>42</sup> Ginet, *Knowledge, Perception, and Memory*, 19.

<sup>43</sup> See also Burstein, "Situating Experience: Agency, Perception, and the Given," 11, where Burstein analyzes Alston's view that an animals' direct access to particulars allows for successful navigation.

<sup>44</sup> See also Alston, "Sellars and the 'Myth of the Given,'" 73.

<sup>45</sup> John McDowell, *Perception as the Capacity for Knowledge* (Milwaukee, WI: Marquette University Press, 2011), 19-21. The language McDowell uses is that the knowledge of the rational and non-rational are two "species" of the same "genus."

I maintain that in the pre-linguistic, the immediate awareness itself provides sufficient warrant for belief, while in the post-linguistic, concepts and background knowledge are required due to how sophisticated the knowledge is.

This view of conceptless belief also accords well with what we know to be the subject's ability to find similarity and discriminate between qualities before propositional knowledge has a chance to be formed. For example, Bayer argues that non-conceptual awareness of similarity between objects is what allows for learning concepts.<sup>46</sup> It is difficult to see how this awareness can be anything but a certain kind of non-conceptual knowledge, for indeed a subject must know there is a similarity before she is able to apply a concept to the similarity. Likewise, J. Adam Carter and Duncan Pritchard point out the importance of discrimination to the gaining the perceptual knowledge: to know that one is in the presence of a zebra, one must discriminate a zebra from non-zebras.<sup>47</sup> The ability to discriminate seems to imply a crude non-conceptual knowledge prior to the formation of conceptual knowledge, for the subject would need to know there are differences in qualities, and what those differences are, in order for there to be a discrimination and a concept to be learned or applied.

### **7.3 Justification**

It is commonly held by those who deny that non-conceptual content exists that non-conceptual content, because it lacks any propositional content, fails to provide any justificatory power to beliefs. Thus, non-conceptual content cannot serve as a foundation for knowledge.<sup>48</sup> This section will be a discussion of how non-conceptual content may confer justification, particularly for propositional beliefs.

---

<sup>46</sup> Bayer, "A Role for Abstractionism in a Direct Realist Foundationalism," 369.

<sup>47</sup> J. Adam Carter and Duncan Pritchard, "Perceptual Knowledge and Relevant Alternatives," *Philosophical Studies* 173 (2016): 969-970. See also Alston, "Sellars and the 'Myth of the Given,'" 71-72.

### 7.3.1 *Recognitional Skill Theory*

One theory of non-conceptual content conferring epistemic justification is that of Stephen L. Reynolds. Reynolds maintains that perceptual beliefs are at least partially justified via a relation with perceptual experiences—experiences being, for Reynolds, non-propositional and non-intentional.<sup>49</sup> This relationship between perception and belief is not one of inference, and cannot be since Reynolds holds that perceptual experiences cannot be true or false.<sup>50</sup> Instead, Reynolds advocates for another kind of justification-conferring normative connection between perception and belief, one that is akin to mastering a practical skill.<sup>51</sup>

Reynolds argues that, when it comes to the practice of a skill, such as playing the piano or speaking a language, we can and do regularly evaluate which performances are good and which performances are unsatisfactory.<sup>52</sup> Moreover, there seem to be rules which people follow in practicing these skills, although belief about following the rules and/or being conscious of the rules during performance doesn't seem necessary for adequate performance, especially at a proficient level.<sup>53</sup> Hence, Reynolds maintains, performing a practical skill is a matter of *knowing-how*, as opposed to *knowing-that*.<sup>54</sup> In the same manner, perceptual knowers have learned how to form certain kinds of appropriate beliefs in response to certain kinds of experiences, that is, they have “recognitional *skills*.”<sup>55</sup>

---

<sup>48</sup> Laurence Bonjour, “Back to Foundationalism,” 70-71.

<sup>49</sup> Steven L. Reynolds, “Knowing How to Believe with Justification,” *Philosophical Studies* 64 (1991): 273.

<sup>50</sup> *Ibid.*, 278-279.

<sup>51</sup> *Ibid.*, 280.

<sup>52</sup> *Ibid.*, 280-281.

<sup>53</sup> *Ibid.*, 281-282.

<sup>54</sup> *Ibid.*, 281.

<sup>55</sup> *Ibid.*, 282.

This comports with our colloquial understanding of some facets of expertise and perception, an understanding which can be extended to perception in general. As Reynolds argues:

Ordinary language admits the existence of skilled judges of music or horses or wines or paintings. Part of what makes them skilled judges is that they are more capable of arriving at justified perceptual beliefs about their subject matter than are others who lack their talent and training. They are more sensitive to the relevant perceptual differences than are people who haven't had similar instruction and practice. But the differences to which they are sensitive are relevantly similar to the features that we become sensitive to in learning to recognize people and objects.<sup>56</sup>

Reynolds argues that if a perceiver's recognitional skills are adequate, the beliefs she forms on the basis of their exercise are justified, that is, barring any undermining beliefs.<sup>57</sup>

Reynolds addresses three objections to his theory, based on how the analogy between practical skills and recognitional skills seems to break down. The first objection is that practical skills can be viewed normatively as rule following—however, there don't seem to conceivably be any rules in the perceptual case. Reynolds counters this by maintaining that recognitional rules are written for robots (what would be better construed today as *artificial intelligence* or *smart devices*), and, further, that these rules would reflect how humans exercise recognitional skills in response to stimuli.<sup>58</sup> Moreover, the fact human beings are able to form an infinite number of beliefs means that “arriving at beliefs must be composed of re-combinable sub-skills, which, apparently, roughly correspond to *some* of the nouns and predicates of our language.”<sup>59</sup> Since acquiring a language is a skill based on rules, so must acquiring perceptual beliefs be a skill based on rules.<sup>60</sup>

---

<sup>56</sup> Ibid.

<sup>57</sup> Ibid., 273, 282.

<sup>58</sup> Ibid., 283-284.

<sup>59</sup> Ibid., 285.

<sup>60</sup> Reynolds speculates that there even may exist “a tenuous connection between truth [i.e., of sentences] and justification by recognitional skills.” However, since he doesn't believe this is important to

The second objection has to do with the performance aspect of practical skills. The objection goes that, since performing a skill is an action that a person does, whereas perception seems like something that just happens automatically (like digestion), then perception is not something that can be normatively evaluated like practical skills. Without this ability to be evaluated, perception is useless for justification purposes.<sup>61</sup> In response, Reynolds points out the performance of practical skills also tends to happen automatically (i.e., reflexively, or without deliberate conscious action), and, in fact, is required to happen automatically when mastered.<sup>62</sup> Automatic performance therefore does not bar a performance from being normatively evaluable. An experienced fencer's responses to his opponent would be automatic and reflexive compared to a novice fencer who must slowly and deliberately arrive at the correct responses.<sup>63</sup>

The last objection Reynolds addresses states that there is no real tie between justification and truth on his view since the correctness or incorrectness of exercising any skill is the product of "arbitrary cultural history."<sup>64</sup> In response, Reynolds entertains the idea that the goal of exercising recognitional skill is arriving at the truth in the same manner as the goal of exercising practical skill is arriving at aesthetic beauty.<sup>65</sup> He admits that this seems implausible, however; subjects do not seem to have the goal of truth in mind when they perceive.<sup>66</sup> Instead, he arrives at an interpretation of recognitional normativity as a

---

answering the above objection, he fails to describe what this might be. See Reynolds, "Knowing How to Believe with Justification," 286.

<sup>61</sup> Reynolds, "Knowing How to Believe with Justification," 286-287.

<sup>62</sup> *Ibid.*, 287.

<sup>63</sup> *Ibid.*

<sup>64</sup> *Ibid.*, 283.

<sup>65</sup> *Ibid.*, 288.

<sup>66</sup> *Ibid.*

“conformity to epistemic norms.”<sup>67</sup> Even so, there is still a question of objectivity, i.e., how justification via recognitional skill relates to truth. In answer, Reynolds argues that the practice of recognitional skill arrives at true beliefs when properly performed because proper performance is affirmed and reinforced by actions having success and public third-party approval.<sup>68</sup> Moreover, experience itself is related to truth in that “experience conveys information of the truth.”<sup>69</sup> However, experiences themselves contain much more information than subjects can perform recognitional skills on, hence different perceptual beliefs about an experiential scene are formed when component information about said experience is processed by the subject.<sup>70</sup>

Reynolds’ view of justification fits well with my view of perception in that it accounts for the distinction between novice and expert recognizers, thus comporting with the idea that the content of perceptual knowledge admits of degrees depending on the subject’s linguistic grasp. The difference between whether a subject is engaging in *perception* and *recognition* would depend on the amount of *recognitional skill* she has, if Reynolds’ view is correct. Moreover, the idea that subjects get better at perception over time, and eventually can do it automatically, fits well with my theory that concepts are learned and slowly applied inferentially before they are automatically applied via *recognition* (see chapter 5, section 5.4). The question becomes then whether or not Reynolds’ view is a plausible account of justification and thus is a good view to adopt.

It seems that, in Reynolds’ view, what does the justificatory work for perceptual beliefs is the successful exercise of recognitional skill itself.<sup>71</sup> It may be argued that this is

---

<sup>67</sup> Ibid.

<sup>68</sup> Ibid., 289-290.

<sup>69</sup> Ibid., 291.

<sup>70</sup> Ibid.

<sup>71</sup> Ibid., 282.

therefore an externalist view of justification (and therefore undesirable), with the exercise of a skill being an external action outside the cognitive purview of a knower, and justification merely relying on success rather than the subject's cognitive grasp. However, Reynolds affirms internalism and seems to maintain that what makes his view internalist is that fact that a subject can reflect on the exercise of this skill and whether or not performance is going well.<sup>72</sup>

Accepting that reflection on performance is an internal mental process, there still is a question of measuring success. This is important to Reynolds' theory since he argues that success of a recognitional skill being exercised reinforces the correctness of performance and ensures that it will be repeated in the future. However, success isn't measured via the ability to get at the truth, but by public feedback.<sup>73</sup> This view of success is open to objections on the basis of relativism, which Reynolds acknowledges.<sup>74</sup> As way of defense, he makes this charge:

If my conjectures are roughly correct (and no doubt they need refinement), then, presumably, accepted epistemic practices which seem to us not to justify the resulting beliefs, fail to do so, in our opinion, because we think that some other pressures guided their adoption. They were adopted for some reason other than success in the resulting actions and the appropriate approval from the community. The dubious epistemic practices of primitive societies are seldom purely recognitional in character...Other societies' adoption of epistemic norms which seem to us not to be justifying will also, on examination, be found to have been wrongly influenced, perhaps by religion or politics, and so to have been improperly learned. The skills approach will thus rule that these practices are not epistemically justifying.<sup>75</sup>

Along with having an ethnocentric standard of justification, Reynold's fails to account for objectivity. How can he claim subjects can be "wrongly" influenced when there's no

---

<sup>72</sup> Ibid., 274, 281-282.

<sup>73</sup> Ibid., 290.

<sup>74</sup> Ibid.

<sup>75</sup> Ibid.

objective standard by which to judge them? Justification, for Reynolds, seems to happen completely independently of truth.

Where Reynolds thinks truth comes in is at the level of experience. This is because, for Reynolds, experience conveys information.<sup>76</sup> However, since the exercise of recognitional skill is not measured in terms of the truth but of success, this explanation still fails to account for how it is that recognitional skill captures accurate information. If experience is infallible, this may work, but then there needs to be an explanation of how a subject can produce a false belief. If it is the case that, in the instance of a false belief, the subject *failed* in the exercising of recognitional skill (i.e., failed to exercise the skill *well*), then arguably the exercise of recognitional skill *does* aim at truth—as opposed to just mere success—and can be *evaluated* in terms of truth.

Reynolds, however, dismisses this view of truth because he doesn't believe perceiving subjects always have truth as a goal. He points out that piano students don't always have beauty as a goal, and so it makes no sense to expect perceivers to have a goal.<sup>77</sup> With this, it seems that Reynolds is backing off the heavily normative and teleological nature of having truth as a goal without even entertaining the possibility. However, appealing to this normativity and teleology is what I think will make Reynolds view most plausible. For even though not every student has the goal of beauty, they *should* have that goal, and in fact part of their instruction is teaching them to *pursue* this goal. Moreover, those formulating beliefs based on perception should have the goal of truth, and teaching children to recognize objects teaches them to pursue this goal in their interactions with the world.

Furthermore, contrary to Reynolds' position, success correlates with truth not because success is measurable by community feedback from other subjects, but because it is

---

<sup>76</sup> Ibid., 291.

<sup>77</sup> Ibid., 288.

measurable by feedback from substances with causal power in the environment. For example, I am walking along the sidewalk and fail to recognize that the sidewalk is unevenly paved. I trip and fall. This causal feedback will teach me to be more attentive to uneven paving surfaces, and what they look like, and avoid tripping in such a manner in the future. Another example is a common toy that children use to learn three dimensional shapes and their relation to two dimensional shapes. There is a wooden box with round, square, and triangular holes, and, further, spherical, cubical, and prismatic blocks that the child is supposed to fit through the holes in order to get them in the box. The feedback the child will receive from doing this exercise incorrectly (she can't put the wrong shape block into the wrong hole) will teach her the nature of the shapes involved. The success of her learning the true nature of these shapes is measured in the success of her overcoming causal obstacles and getting the shapes into the box. This better captures the consequential nature of objective truth than Reynolds' theory about community feedback.

### ***7.3.2 Justification Through Perceptual Content***

Understandably, some may not want to accept a theory of justification that relies so heavily on a such a normative and teleological conception of the subject's perceptual goals (i.e., those who may be put off by teleological explanations in general). As such, it may be beneficial to take a more intuitive route and advocate for a view of justification where the immediate content of perception itself does justificatory work. This notion of justification was already hinted at above in section 7.2 in the discussion of how pre-linguistic belief is warranted compared to post-linguistic. I will now develop this view further as a theory of justification via an analysis of the work of Carl Ginet and William Alston.

Like the view being developed in this thesis, Ginet argues that knowledge is not always accompanied by the subject's ability to express knowledge.<sup>78</sup> Even some complex

---

<sup>78</sup> Ginet, *Knowledge, Perception, and Memory*, 25-26.

recognitional knowledge needs not be expressible by post-linguistic subjects in order for it to be considered knowledge:

And we ourselves know many things that we would be at a loss to put into words...I know what the characteristic appearance of a certain person is that one may go by in recognizing him, as I show by recognizing him, or pictures of him, very well; but if I am asked to *say* just what that characteristic appearance is I am able only to point to him, or a picture of him, and say, 'Well, it's what can be seen there' (though this *might* be fairly regarded as *saying* what he looks like).<sup>79</sup>

Thus, Ginet's view captures the complexity of my view of *recognitional* knowledge as well as its non-inferential nature. For indeed, if one was to *infer* their way to recognizing a face, they should be able to articulate the various features which allowed this inference.

Ginet also ties the ability to recognize to concept deployment, arguing that a subject will be able to perceptually recognize—and therefore know—that certain facts obtain so long as he possesses the relevant sort of concepts and is able to reflect on whether the facts do obtain.<sup>80</sup> Additionally, Ginet points out that multiple propositional descriptions can fit a singular visual scene, for example, that a puddle of water under normal conditions looks exactly like a puddle of gasoline.<sup>81</sup> This comports with my view that multiple concepts may be justifiably applied to the same perceptual experience, that the contents of knowledge may vary between subjects depending on how many or which accurate and relevant concepts have been applied, and that concepts can be justifiably but wrongly applied even by the knowledgeable. Thus, even amongst subjects with comparable linguistic skills and background knowledge, there may be disagreement. This would explain how objective moral properties can produce the same qualitative experience across subjects and yet would still result in wildly varying moral judgements between subjects.

---

<sup>79</sup> Ibid., 25.

<sup>80</sup> Ibid., 34.

<sup>81</sup> Ibid., 99-101.

Justification for belief comes by way of these recognized facts directly influencing a subject's doxastic attitude. As Ginet explains:

A fact can be part of what justifies *S* in being confident that *p* only if it is a fact that can *directly* influence *S*'s doxastic attitude towards that proposition. That is, *S*'s doxastic attitude at any given time towards any given proposition can be justified or unjustified only on the basis of what at that time requires only *S*'s effort of attention or consideration in order to influence his attitude.<sup>82</sup>

Only those items that can influence his doxastic attitude which are accessible to the subject are relevant for justificatory purposes.<sup>83</sup> Ginet also adds that, in order for justification to hold, the subject must lack a reason not to trust his perception.<sup>84</sup>

However, it is unclear on Ginet's view whether or not what does the justificatory work for a belief also needs to be known to the agent as a justifier. As Ginet states:

The requirement of direct recognizability on justification for confidence (or justification for any other degree of belief)—that is, the requirement that any minimally sufficient condition for *S*'s having justification for being confident that *p* be directly recognizable to *S*—can be seen to hold by the following argument.

Assuming that *S* has the concept of justification for being confident that *p*, *S* *ought* always to possess [sic] or lack confidence that *p* according to whether or not he has such justification. At least he ought always to withhold confidence unless he has justification. This is simply what is meant by having or lacking *justification*. But if this is what *S* ought to do in any possible circumstance, then it is what *S* *can* do in any possible circumstance. That is, assuming that he has the relevant concepts, *S* can always tell whether or not he has justification for being confident that *p*.<sup>85</sup>

Thus, on one hand, Ginet argues that recognized facts influence a subject's doxastic attitude, and thereby justify a subject's belief. On the other hand, as this quote above states, in order for the subject to be justified in his belief, he needs to know that he is justified and recognize what makes him justified. This seems part of his justification, at least, if he has the concept of justification.

---

<sup>82</sup> Ibid., 34.

<sup>83</sup> Ibid.

<sup>84</sup> Ibid.

<sup>85</sup> Ibid., 36.

It is unclear, however, what happens to subjects who lack a concept for justification, like the pre-linguistic. Arguably, these subjects can be justified merely by the (non-conceptual, non-propositional) features of the immediate visual scene. This seems to fit with Ginet's earlier view that the perceived having a direct influence on doxastic attitudes is what confers justification. However, pre-linguistic subjects might not know that or how a scene confers justification given that they lack a concept for justification, and therefore their beliefs may not be ultimately justified under Ginet's view reflected in the latter quote. The question becomes, do subjects need to know they are justified—and need to know what makes them justified—in order for them to be justified? To ask this question is not to cast doubt on internalism when it comes to justification. It is to ask whether the subject's internal justifying state is the non-conceptual content—and any relevantly applied conceptual content—*itself*, or if it includes the subject's awareness *that* the content justifies and *how* the content justifies.

As Alston points out, the claim that the subject needs to know the epistemic status of her belief—in this case, that and why her belief is justified—can “be used against *any* claim to immediate knowledge.”<sup>86</sup> It isn't a fair defeat of immediate knowledge, however; demanding that the subject be self-reflective in order to be justified is, as Alston puts it, “confus[ing] epistemology with its own subject matter.”<sup>87</sup> Alston argues that this problem comes by way of equivocating between being justified in one's belief and justifying one's belief, and believing that both need to obtain for a belief to be justified.<sup>88</sup> However, for a belief to be justified, only the former needs to obtain. The latter is merely a communicable feature of justification possessed by post-linguistic subjects who possess the concept of justification. A subject, however, can be completely justified in her belief without knowing

---

<sup>86</sup> Alston, “What's Wrong with Immediate Knowledge?” 81.

<sup>87</sup> *Ibid.*, 86.

<sup>88</sup> Alston, “Sellars and the ‘Myth of the Given,’” 85.

that she is justified or the reason why she is justified.<sup>89</sup> So, just as a pre-linguistic subject can possess a belief without being able to express the belief, a pre-linguistic subject can also be justified without being able to express her justification.

What does justification consist of? In other words, how are perceptual beliefs justified? Alston argues that perceptual beliefs are justified because they are reflective of the perceptual experience subjects undergo:

We are able to justifiably form beliefs about the external environment on the basis of our perceptual experience because objects in the external environment appear to us in that experience in such a way as to be constitutive of the character of the experience. And the beliefs so formed are *prima facie* justified just because they *register* what is presented there, they “read it off of” experience, possibly corrected in the light of whatever independent knowledge we bring to bear.<sup>90</sup>

According to Alston, the subject’s belief is either justified by the experience itself or because it arises from the experience. There is no need for a subject to know that she is undergoing a justificatory experience. It would be unreasonable to require that the subject has this kind of higher order knowledge anyway, since it is typical for subjects to solely attend to the content of their experiences and not to the having of experiences.<sup>91</sup>

Alston also confirms that background knowledge does play a role in a perceptual belief being justified. The example he utilizes to demonstrate this is of going to someone’s house that we’ve gone to before. When looking for a house that we’ve been to before, especially on numerous occasions, we will recognize it by the way it appears. This, in addition to the (not necessarily conscious) knowledge that we are in the right neighborhood, etc., will justify us in believing we’ve found the right house when we see it.<sup>92</sup>

---

<sup>89</sup> Alston, “What’s Wrong with Immediate Knowledge?” 75, 83-86, 90.

<sup>90</sup> Alston, “Back to the Theory of Appearing,” 198.

<sup>91</sup> Alston, “Perceptual Knowledge,” 227.

<sup>92</sup> *Ibid.*, 228.

Alston's view works well with the view of perceptual knowledge being developed here since it coheres with the notion that post-linguistic concept holders' knowledge differs in content than the pre-linguistic, and, further, that the post-linguistic are held to different justificatory standards than the pre-linguistic considering the amount of background knowledge and concepts they possess. In other words, post-linguistic subjects are held epistemically accountable to the degree of background knowledge (including concepts) that they possess when they engage in an instance of forming perceptual knowledge. Since the beliefs that result from their perceptions are more complex, they require more justification in order to be justified.

Contrary to the view being propagated here, however, Alston assumes that perceptual knowledge always has a conceptual element to it. Thus, on this basis he ignores the question of whether conceptualization is essential to perception.<sup>93</sup> As already argued for in section 7.2, however, perceptual knowledge need not have a conceptual element in the pre-linguistic as long as we accept the knowledge they hold is very limited, possibly confined to the basic discrimination of qualities. For example, a young child may have knowledge of a particular quality she sees in a visual scene because she is able to discriminate it from other qualities. Hence, she can differentiate between green and red and have knowledge of the quality of each without possessing a concept for either.

Also integral to perceptual justification is the absence of defeaters, including suspicions of abnormal perception and strong counter beliefs.<sup>94</sup> If we have sufficient reason to doubt what we perceive to be true, then our belief loses justification. But as Alston points out, it isn't that normal perception confers justification so that we need to be aware that our perceptions are normal in order to be justified. As Alston puts it, "normality considerations

---

<sup>93</sup> Ibid., 232.

<sup>94</sup> Ibid., 228; Reynolds, "Knowing How to Believe with Justification," 273. See also Michael Huemer, "Compassionate Phenomenal Conservatism," *Philosophy and Phenomenological Research* 74 (2007): 30.

can figure as possible overrides of *prima facie* justification, rather than as part of what confers *prima facie* justification.”<sup>95</sup> The subject still can be justified in her belief without any higher-order awareness that her experience is normal if it isn’t abnormal. Thus, without sufficient reason to suspect otherwise, we are justified in taking our perceptions as true.

A common objection to this view of justification is to point to hallucinations. If a hallucinating subject has no reason to suspect that his experiences are not veridical, then he would be justified in holding false beliefs under the view being advanced here. Moreover, there would be no reason to hold beliefs as justified under normal conditions since there’s no way of telling whether the experience is real or hallucinatory. The same goes with dreaming or any sensory state that occurs without external cause.<sup>96</sup> Since we can never know whether our perceptions are veridical, our perceptions cannot justify our perceptual beliefs.

Alston has two answers to this type of objection. The first is to say that the subject’s belief in the case of a hallucination is true, but it is a belief about a mental image, not anything external. Hence, the only thing I could be justified in is my belief about my own internal states. The second is to say experiences only justify that which is directly accessible to the subject, and given that whether or not an experience is real or hallucinatory is not accessible, what is justified is neutral in terms of whether it is real or hallucinatory.<sup>97</sup> Thus, I can believe that there is an oak tree outside my window, and be justified in that belief, but this does not give credence to the idea that an oak tree is anything external to me. If these are the only solutions, then the theory here is undesirable, especially since what we are hoping to capture as true are states of affairs in the *external* world.

---

<sup>95</sup> Alston, “Perceptual Knowledge,” 228-229.

<sup>96</sup> See Ernest Sosa, “Skepticism and the Internal/External Divide,” in *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa (Malden, MA: Blackwell Publishing, 2003), 142.

<sup>97</sup> Alston, “Perceptual Knowledge,” 239-240.

Alston's solution to this problem is to affirm that the contents of experience, and their ability to justify, *is* neutral in terms of whether the perception is real or hallucinatory, but that beliefs about anything directly accessible to the subject are *prima facie* justified unless the subject herself is justified in believing that what she is experiencing is a hallucination.<sup>98</sup> Hence, I am justified directly through experience in my belief about an oak tree outside my window, that it is a real external object, so long as I am not justified in believing I am hallucinating.

I go a little further than this and argue that the external universe is more ordered and coherent than dreams and most hallucinations, and in obvious ways. Thus, the disordered nature and incoherence of false perceptions should be sufficient reason for any beliefs based upon them to be unjustified. Concerning forming beliefs on the basis of hallucinations that mimic reality in order and rationality, I take the subject to have formed a justified belief—just not one that is true. Thus, in this case of a false justified belief, the subject lacks knowledge.

I don't think the concession of allowing for justified false beliefs works against the theory being put forward here, unless one wants to adopt extreme skepticism. For I don't believe the existence of reality-mimicking hallucinations gives us a reason to doubt all of our perceptions, so that none can be trusted. Our causal interactions with the external world support the view that our perceptual capacity is, on the whole, reliable. If it wasn't reliable, we would not be able to successfully navigate our environment. This reliability gives us justification to trust our senses. Therefore, the existence of hallucinations does not entail that perceptual experiences cannot directly justify perceptual beliefs.

---

<sup>98</sup> *Ibid.*, 240.

Another objection along these same lines is that we are not always correct concerning what we perceive; we can be mistaken in our perceptions. Thus, perception cannot be immediate nor our perceptual beliefs directly justified by perception.<sup>99</sup> However, like Alston, I affirm that the immediacy of perception doesn't entail that perceptual beliefs are infallible.<sup>100</sup> As Alston explains, immediacy has to do with the lack of mediation in awareness, not with unshakable assurance that an object is exactly as it appears or exactly as we interpret its appearance.<sup>101</sup> Moreover, we know that appearances can be misleading, and if appearances were always or mostly misleading, then we would have reason to question the epistemic value of perception.<sup>102</sup> But we know from our causal contact with the external world that our perceptions are typically reliable,<sup>103</sup> so it is reasonable to conclude that our perceptions aren't as on such shaky epistemic ground as this objection implies.

#### **7.4 Non-Conceptual Content and Justified Moral Perceptual Beliefs**

So where does this discussion on non-conceptual content and justified perceptual beliefs leave us? Recall that in our theory of moral perception we wanted to make the case that phenomenal experience of moral properties was the same across agents regardless of their moral maturity, but where the morally immature could merely *perceive* moral phenomena, the morally mature could *recognize* moral phenomena. The difference between these—*perception* and *recognition*—was the deployment of relevant moral concepts *on sight* (i.e., without the aid of inference) by the subject: the recognizer possesses the relevant moral concepts, while the mere perceiver does not.

---

<sup>99</sup> Alston, "What's Wrong with Immediate Knowledge?" 77.

<sup>100</sup> Ibid.

<sup>101</sup> Alston, "Back to the Theory of Appearing," 183.

<sup>102</sup> Ibid.; Alston, "Perceptual Knowledge," 233.

<sup>103</sup> Alston, "Sellars and the 'Myth of the Given,'" 73.

This inability to deploy relevant concepts does not hinder the formation of justified belief. That is, in both cases of perception and recognition,<sup>104</sup> perceptual knowledge is gained. However, what the inability to deploy the relevant concepts does hinder is the *amount* of moral perceptual knowledge that can be gained given a moral scene. Hence, perceptual knowledge across subjects varies in degrees depending on the background knowledge, including the concepts, possessed by each subject. For the morally immature who do not possess any relevant moral concepts, knowledge may only be limited to the moral qualities produced by the moral scene. For example, just like the young child whose perceptual knowledge of the color *red* only consists in knowing the quality of red and being able to differentiate it from the quality of green, the morally immature may only know what the quality of a moral scene is like (in a what-it-is-like sense) and be able to discriminate it from other qualities. In this case the morally immature are only able to engage in *perception* and not in any *recognition*.

A slightly more morally mature subject who possesses the moral concept *bad* may recognize that the moral quality of the scene is *bad*. She would therefore be justified in knowing that it is *bad*, but still may lack the recourse to know why it is *bad*. A more morally mature person, on the other hand, may have knowledge of what the moral quality is (e.g., it is a *torture* of a cat) and be able to employ all the background knowledge that the deployment of a relevant moral concept (in this case *torture*) carries into the scene. Thus, the more morally mature a subject is, the fuller her perceptual knowledge will be given the same perceptual scene. For the subject who possesses no moral concepts, their knowledge is justified completely through the immediate awareness of the moral phenomena. For the more

---

<sup>104</sup> Recall from chapter 5 that I explain that *recognition* is a type of perception, i.e., all recognition is perception, but not all perception is recognition. See chapter 5, section 5.4.

mature, it is justified both by immediate awareness and through background conceptual knowledge.

Thus, making sense of moral perceptual knowledge in this fashion allows an explanation for why different moral agents sometimes come to different moral judgments concerning the same moral scene even given such a robust phenomenal moral realism: different agents may possess full understanding of the concept (or concepts) necessary for complete recognition while others may only have partial understanding of a moral concept and thus only achieve partial recognition.<sup>105</sup> For example, where someone may only recognize that an act is *bad*, another may recognize that an act is *torture*. The subject who only recognizes the *badness* of the moral scene may not understand why it is bad, only that it should not be happening. The subject who recognizes *torture* understands that the concept *torture*—along with it being *bad*—carries with it concern about sentient life and how it should and should not be treated. Thus, in the instance of coming across a scene of youths torturing a cat, both the subject who possess the concept *torture* and the subject who lacks it will gain perceptual moral knowledge of the scene. However, the content of that knowledge will differ for each.

We also have an explanation for why virtuous agents have the ability to come to complex moral judgments so quickly given a moral situation: they have mastered and learned to deploy *on sight* many different kinds of complex moral concepts. With this, we can see why a more Reynoldsian approach to moral perception may also be fitting: in perceptual moral instances where the more morally immature would need to employ practical reasoning (i.e., methodically apply what they are learning to the moral situation at hand), a morally virtuous agent would be able to recognize immediately on sight the true nature of the moral

---

<sup>105</sup> This does not bar that instances of perceptual moral disagreement can and do stem from some agents just being wrong about what they see or deploying completely inappropriate concepts to the situation.

situation, reflexively come to the correct moral judgement, and spring reflexively into a fitting moral action. Because the virtuous agent has mastered such a moral recognitional skill, she would be, if we took a Reynoldsian approach, justified in her perceptual beliefs. In either case, whether Reynoldsian or Alstonian, the virtuous person is justified immediately and non-inferentially.

## Conclusion

I had made the claim in the thesis introduction that the purpose of this thesis was to reconnect ethics and metaphysics in a serious way, that is, in a way that post-enlightenment moral philosophers largely fail to do. I also characterized this thesis as defending the notion that moral properties are a certain class of ordinary properties, and moreover, that they have their own sense-perceptible phenomenology. What I hope is now apparent is that human beings live in an inherently moral world which entails that they regularly commit moral actions and constantly possess moral obligations. Human beings cannot help but to live in this reality.

For one, we live in a world where substance causation holds true (chapter 1) and final causes bestow upon substances a normative causal trajectory (chapter 6). Human beings are substances with causal powers and free originators of causal chains (chapters 1, 2, and 4). Due to their causal abilities coupled with knowledge of the causal tendencies and normative trajectories of substances, human beings are morally culpable beings (chapter 6). Since human beings cannot fail to causally participate in an inherently causal universe, they cannot avoid moral action. This even includes those instances where human beings fail to act since their causal capabilities and knowledge could have resulted in a more morally desirable causal chain of events. This also includes instances where no causal activity is happening, since causally salient relations between moral agents and substances still exist and therefore moral obligations still apply (chapter 6).

Second, the moral properties that ground this moral realism are sense-perceptible. Like ordinary natural properties (and because they are ordinary natural properties), they have their own what-it-is-like (chapters 3, 5, and 6). Moral properties have this because they are *relational* properties; they are not *sui-generis* or *supervenient*, which would bar them from having their own externally generated phenomenology (chapter 5). Specifically, moral

properties are relational *features* of relations, i.e., the total complex states of affairs between one or more substances (chapter 6). Moreover, the perception of moral qualities is experienced by all normally functioning subjects and is enough for the formation of perceptual moral knowledge (chapters 5, 6, and 7). Differences in the content of moral perceptual knowledge between subjects can be accounted for through the background knowledge and moral concepts that subjects possess (chapter 5 and 7).

We can see now that the two oughts from the thesis introduction—of *something ought to be done* and *the agent ought to do something about it*—is grounded in the very fabric of the universe. The imperative arises via the nature of human beings—given their knowledge and causal capability—and the nature of reality—being so heavily causally constructed—that human beings just are moral creatures in a moral world. They are morally acting and acted upon daily, and constantly are in relations that entail moral obligations. This is the nature of human beings in our universe: moral agents in a moral world.

## Bibliography

- Adams, Robert M. "Idealism Vindicated." In *Persons: Human and Divine*, edited by Peter van Inwagen and Dean Zimmerman, 35-54. Oxford: Oxford University Press, 2007.
- Adams, Russell J. "An Evaluation of Color Preference in Early Infancy," *Infant Behavior and Development* 10 (1987): 143-150.
- Alston, William P. "An Internalist Externalism." *Synthese* 74 (1988): 265-283.
- , "Back to the Theory of Appearing." *Philosophical Perspectives* 13 (1999): 181-203.
- , "Level Confusions in Epistemology." In *Epistemic Justification: Essays in the Theory of Knowledge*, 153-171. Ithaca, NY: Cornell University Press, 1989.
- , "Perceptual Knowledge." In *The Blackwell Guide to Epistemology*, edited by John Greco and Ernest Sosa, 223-242. Malden, MA: Blackwell Publishers, 1999.
- , "Sellars and the 'Myth of the Given.'" *Philosophy and Phenomenological Research* 65 (2002): 69-86.
- , "The Deontological Conception of Epistemic Justification." *Philosophical Perspectives* 2 (1988): 257-299.
- , "What's Wrong with Immediate Knowledge?" *Synthese* 55 (1983): 73-95.
- Alvarez, Maria. *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press, 2010.
- Armstrong, D.M. *Sketch for a Systematic Metaphysics*. Oxford: Oxford University Press, 2010.
- Audi, Robert. *Epistemology: A Contemporary Introduction to the Theory of Knowledge, 3<sup>rd</sup> Edition*. New York: Routledge, 2011.
- , "Moral Knowledge and Moral Perception." *Proceedings of the Aristotelian Society, Supplementary Volumes* 84 (2010): 79-97.
- , *Moral Perception*. Princeton, NJ: Princeton University Press, 2013.
- Austin, Christopher J. "Is Dispositional Causation Just Mutual Manifestation?" *Ratio* 29 (2016): 235-248.
- Averill, Edward Wilson. "The Phenomenological Character of Color Perception." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 157 (2012): 27-45.

- Barnett, David. "You are Simple." In *The Waning of Materialism*, edited by Robert C. Koons and George Bealer, 161-174. Oxford: Oxford University Press, 2010.
- Baumgartner, Michael. "Regularity Theories Reassessed." *Philosophia* 20 (2008): 327-354.
- Bayer, Benjamin. "A Role for Abstractionism in a Direct Realism Foundationalism." *Synthese* 180 (2011): 357-389.
- Berkeley, George. *A Treatise Concerning the Principles of Human Knowledge and Three Dialogues Between Hylas and Philonous*, edited by Roger Woolhouse. London: Penguin Books Ltd, 1988.
- Beverluis, John. "Some Remarks Concerning a Non-propositional Knowledge of God." *Australasian Journal of Philosophy* 43 (1965): 376-381.
- Bird, Alexander. "Fundamental Powers, Evolved Powers, and Mental Powers." *Aristotelian Society Supplementary* 92 (2018): 247-275.
- , "Overpowering: How the Powers Ontology Has Overreached Itself." *Mind* 125 (2016): 341-383.
- , "The Regress of Pure Power?" *The Philosophical Quarterly* 57 (2007): 513-534.
- Bonjour, Laurence. "Back to Foundationalism." In *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa, 60-76. Malden, MA: Blackwell Publishing, 2003.
- , "The Regress Problem and Foundationalism." In *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa, 5-23. Malden, MA: Blackwell Publishing, 2003.
- Brown, Harold I. "The Case for Indirect Realism." In *The Case for Qualia*, edited by Edmond Wright, 45-58. Cambridge, MA: MIT Press, 2008.
- Buras, Todd and Trent Dougherty. "Parrying Parity: A Reply to a Reidian Critique of Idealism." In *Idealism: New Essays in Metaphysics*, edited by Tyron Goldschmidt and Kenneth L. Pearce, 1-16. Oxford: Oxford University Press, 2017.
- Burstein, Matthew. "Situating Experience: Agency, Perception, and the Given." *The Southern Journal of Philosophy* 44 (2006): 1-29.
- Carter, J. Adam and Duncan Pritchard. "Perceptual Knowledge and Relevant Alternatives." *Philosophical Studies* 173 (2016): 969-990.
- Chalmers, David J. "The Combination Problem for Panpsychism." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Bruntrup and Ludwig Jaskolla, 179-214. Oxford: Oxford University Press, 2016.
- , *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press, 1996.

- , "Facing Up to the Problem of Consciousness." *Journal of Consciousness Studies* 2 (1995): 200-219.
- , "Panpsychism and Protopanpsychism." *The Amherst Lecture in Philosophy* 8 (2013): 1-35.
- Churchland, Paul M. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press, 1989.
- Connell, Richard J. *Substance and Modern Science*. Houston, TX: Center for Thomistic Studies, 1988.
- Crane, Tim. "The Nonconceptual Content of Experience." In *The Contents of Experience: Essays on Perception*, edited by Tim Crane, 136-157. Cambridge: Cambridge University Press, 1992.
- Crooks, Mark. "The Churchlands' War on Qualia." In *The Case for Qualia*, edited by Edmond Wright, 203-222. Cambridge, MA: MIT Press, 2008.
- Dancy, Jonathan. *Practical Reality*. Oxford: Oxford University Press, 2000.
- Davidson, Donald. "Causal Relations." *The Journal of Philosophy* 64 (1967): 691-703.
- Deery, Oisín, Matt Bedke, and Shaun Nichols. "Phenomenal Abilities: Incompatibilism and the Experience of Agency." In *Oxford Studies in Agency and Responsibility, Volume 1*, edited by David Shoemaker, 126-150. Oxford: Oxford University Press, 2013.
- Dennett, Daniel C. *Consciousness Explained*. London: Penguin Books Ltd, 1991.
- Dretske, Fred. "Phenomenal Externalism or If Meanings Ain't in the Head, Where Are Qualia?" *Philosophical Issues* 7 (1996): 143-158.
- Ellis, Ralph D. "Three Arguments Against Causal Indeterminacy." *Philosophia* 31 (2004): 331-344.
- Engelhard, Kristina. "Categories and the Ontology of Powers: A Vindication of the Identity Theory of Properties." In *The Metaphysics of Powers: Their Grounding and their Manifestations*, edited by Anna Marmodoro, 41-57. New York: Routledge, 2010.
- Feser, Edward. *Scholastic Metaphysics: A Contemporary Introduction*. Piscataway, NJ: Transaction Books, 2014.
- Flage, Daniel E. "Berkeley's Archetypes," *Hermathena* 171 (2001): 7-31.
- Fine, Kit. "Neutral Relations." *The Philosophical Review* 109 (2000): 1-33.
- Fodor, Jerry. "The Revenge of the Given." In *Contemporary Debates in Philosophy of Mind*, edited by Brian P. McLaughlin and Johnathan Cohen, 105-116. Malden, MA: Blackwell Publishing Ltd, 2007.

- Foster, John. *The Nature of Perception*. Oxford: Oxford University Press, 2000.
- Fulton, Ann, et al. "Development of Primate Rod Structure and Function." In *Infant Vision*, edited by François Vital-Durand, Janette Atkinson, and Oliver J. Braddick, 35-50. Oxford: Oxford University Press, 1996.
- Gibb, S.C. "Mental Causation and Double Prevention." In *Mental Causation and Ontology*, edited by S.C. Gibb, E. J. Lowe, and R.D. Ingthorsson, 193-214. Oxford: Oxford University Press, 2013.
- Ginet, Carl. *Knowledge, Perception, and Memory*. Dordrecht, Holland: Reidel Publishing Company, 1975.
- Graham, George and Terence Horgan. "Qualia Realism, Its Phenomenal Contents and Discontents." In *The Case for Qualia*, edited by Edmond Wright, 89-108. Cambridge, MA: MIT Press, 2008.
- Grzankowski, Alex. "Pictures Have Propositional Content." *Review of Philosophy and Psychology* 6 (2015): 151-163.
- Handfield, Toby. "Dispositions, Manifestations, and Causal Structure." In *The Metaphysics of Powers: Their Grounding and their Manifestations*, edited by Anna Marmodoro, 106-132. New York: Routledge, 2010.
- Harman, Gilbert. *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press, 1977.
- Heil, John. "Powerful Qualities." In *The Metaphysics of Powers: Their Grounding and their Manifestations*, edited by Anna Marmodoro, 58-72. New York: Routledge, 2010.
- Huemer, Michael. "Compassionate Phenomenal Conservatism." *Philosophy and Phenomenological Research* 74 (2007): 30-55.
- Hume, David. *A Treatise of Human Nature*, edited L.A. Selby-Bigge. Oxford: Clarendon Press, 1896.
- Kane, Robert. *A Contemporary Introduction to Free Will*. Oxford: Oxford University Press, 2005.
- Kastrup, Bernardo. "Conflating Abstraction with Empirical Observation: The False Mind-Matter Dichotomy." *Constructivist Foundations* 13 (2018): 341-361.
- , "On the Plausibility of Idealism: Refuting Criticisms." *Disputatio* 9 (2017): 13-34.
- Kim, Jaegwon. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press, 1998.
- Kind, Amy. "How to Believe in Qualia." In *The Case for Qualia*, edited by Edmond Wright, 285-298. Cambridge, MA: MIT Press, 2008.

- Knowles, Robert. "Heavy Duty Platonism." *Erkenntnis* 80 (2015): 1255-1270.
- Koons, Robert C. "Dual Agency: A Thomistic Account of Providence and Human Freedom," *Philosophia Christi* 4 (2002): 397-411.
- Kripke, Saul. *Naming and Necessity*. Malden, MA: Blackwell Publishing, 1981.
- Lewis, David. "Causation." *The Journal of Philosophy* 70 (1973): 556-567.
- , "Causation As Influence." *The Journal of Philosophy* 97 (2000): 182-197.
- Lowe, E. J. "On the Individuation of Powers." In *The Metaphysics of Powers: Their Grounding and their Manifestations*, edited by Anna Marmodoro, 8-26. New York: Routledge, 2010.
- , "Personal Agency." *Royal Institute of Philosophy Supplement* 53 (2003): 211–227.
- , *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford University Press, 2008.
- , "Substance Causation, Powers, and Human Agency." In *Mental Causation and Ontology*, edited by S.C. Gibb, E. J. Lowe, and R.D. Ingthorsson, 153-172. Oxford: Oxford University Press, 2013.
- , "There Are (Probably) No Relations," *The Metaphysics of Relations*, edited by Anna Marmodoro and David Yates, 100-112. Oxford: Oxford University Press, 2016.
- , "The Will as a Rational Free Power." In *Powers and Capacities in Philosophy: The New Aristotelianism*, edited by Ruth Groff and John Greco, 172-185. New York: Routledge, 2013.
- MacBride, Fraser. "How Involved Do You Want to Be in a Non-Symmetric Relationship?" *Australasian Journal of Philosophy* 92 (2014): 1-16.
- Marmodoro, Anna and David Yates. "Introduction." In *The Metaphysics of Relations*, edited by Anna Marmodoro and David Yates, 1-18. Oxford: Oxford University Press, 2016.
- Maslen, Cei. "Pragmatic Explanations of the Proportionality Constraint on Causation." In *Making a Difference: Essays on the Philosophy of Causation*, edited by Helen Beebe, Christopher Hitchcock, and Huw Price, 57-71. Oxford: Oxford University Press, 2017.
- Maund, Barry. "A Defense of Qualia in the Strong Sense." In *The Case for Qualia*, edited by Edmond Wright, 269-284. Cambridge, MA: MIT Press, 2008.
- McDonnell, Neil. "Events and their Counterparts." *Philosophical Studies* 173 (2016): 1291-1308.
- McDowell, John. *Perception as the Capacity for Knowledge* (Milwaukee, WI: Marquette University Press, 2011).

- Molnar, George. *Powers: A Study in Metaphysics*. Edited by Stephen Mumford. Oxford: Oxford University Press, 2003.
- Nida-Rümelin, Martine. "Phenomenal Character and the Transparency of Experience." In *The Case for Qualia*, edited by Edmond Wright, 309-324. Cambridge, MA: MIT Press, 2008.
- O'Connor, Timothy. *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford University Press, 2002.
- O'Dea, John. "Transparency and the Unity of Experience." In *The Case for Qualia*, edited by Edmond Wright, 299-308. Cambridge, MA: MIT Press, 2008.
- Oderberg, David S. "Divine Premotion." *International Journal for Philosophy of Religion* 79 (2016): 207-222.
- , "Finality Revived: Powers and Intentionality." *Synthese* 194 (2017): 2387-2425.
- , *The Metaphysics of Good and Evil*. New York: Routledge, 2020.
- Owen, Matthew. "Aristotelian Causation and Neural Correlates of Consciousness." *Topoi* (2018). <https://doi.org/10.1007/s11245-018-9606-9>.
- Raffman, Diana. "From the Looks of Things: The Explanatory Failure of Representationalism." In *The Case for Qualia*, edited by Edmond Wright, 325-240. Cambridge, MA: MIT Press, 2008.
- Raz, Joseph. *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press, 2002.
- Reynolds, Stephen L. "Knowing How to Believe with Justification." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 64 (1991): 273-292.
- Robinson, Howard. "Why Frank Should Not Have Jilted Mary." In *The Case for Qualia*, edited by Edmond Wright, 123-246. Cambridge, MA: MIT Press, 2008.
- Schellenberg, Susanna. "Perceptual Content Defended." *Nous* 45, 4 (2011): 714-750.
- Shafer-Landau, Russ. *Moral Realism: A Defense*. Oxford: Oxford University Press, 2003.
- Sinclair, Neil. "Reasons Internalism and the Function of Normative Reasons," *Dialectica* 71 (2017): 209-229.
- Sireteanu, Ruxandra. "Development of the Visual Field: Results from Human and Animal Studies." In *Infant Vision*, edited by François Vital-Durand, Janette Atkinson, and Oliver J. Braddick, 18-28. Oxford: Oxford University Press, 1996.
- Smith, Michael. "Internal Reasons." *Philosophy and Phenomenological Research* 55 (1995): 109-131.

- Sosa, Ernest. "Does Knowledge Have Foundations?" In *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa, 119-140. Malden, MA: Blackwell Publishing, 2003.
- , "Skepticism and the Internal/External Divide." in *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence Bonjour and Ernest Sosa, 141-155. Malden, MA: Blackwell Publishing, 2003.
- Stoljar, Daniel. *Physicalism*. New York: Routledge, 2010.
- Strawson, Galen. *Mental Reality*, 2<sup>nd</sup> ed. Cambridge, MA: MIT Press, 2009.
- , "Mind and Being: The Primacy of Panpsychism." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Bruntrup and Ludwig Jaskolla, 75-112. Oxford: Oxford University Press, 2016.
- , "Realistic Monism: Why Physicalism Entails Panpsychism." In *Real Materialism: and Other Essays*, 53-74. Oxford: Oxford University Press, 2008.
- , "Real Materialism." In *Real Materialism: and Other Essays*, 19-51. Oxford: Oxford University Press, 2008.
- Tooley, Michael. "Probability and Causation." In *Cause and Chance: Causation in an Indeterministic World*, edited by Phil Dowe and Paul Noordhof, 77-119. New York: Routledge, 2004.
- Williams, Bernard. "Internal and External Reasons." In *Moral Luck*. Cambridge: Cambridge University Press, 1981.
- Wright, Edmond. "Why Transparency is Unethical." In *The Case for Qualia*, edited by Edmond Wright, 341-366. Cambridge, MA: MIT Press, 2008.